

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1327	US-09-762-277A-1	1 MNPLLILAFGLRAAVATPTDD.....VYTKVCNFVDWIQSTIAANS	247			P06872 canis familiaris
2	1153	BLOSUM62 Gapop 10.0 , Gapext 0.5	112892 seqs, 41476328 residues	100			P29463 bos taurus
3	1127	Searched: ,tal number of hits satisfying chosen parameters:	112892	0			P07146 mus musculus
4	1111	Maximum DB seq length: 0					P00763 rattus norvegicus
5	1101	Maximum DB seq length: 2000000000					P0762 rattus norvegicus
6	1098.5	Post-processing: Minimum Match 0%					P08426 rattus norvegicus
7	1097	Maximum Match 100%					P07478 homo sapiens
8	1085	Listing first 45 summaries					P07477 homo sapiens
9	1071	Database : Swissprot_40.4*					P0761 sus scrofa
10	1058						P15951 homo sapiens
11	1036						P35030 homo sapiens
12	1033						P70059 xenopus laevis
13	1026						P06871 canis familiaris
14	1019						P07660 bos taurus
15	992.5						P19799 xenopus laevis
16	966.5						Q90629 gallus gallus
17	966.5						P12788 rattus norvegicus
18	948.5						P35033 salmo salar
19	940						P0764 squalus acanthias
20	937						Q91041 gadus morhua
21	905						P16049 pierreoneectes
22	896						P35034 pierreoneectes
23	877						P35031 salmo salar
24	867						P35032 salmo salar
25	85.5						P07642 squalius acanthias
26	83.4.5						Q91041.1 gadus morhua
27	82.4.5						P16049.1 pierreoneectes
28	64.9.5						P35034.1 pierreoneectes
29	61.4						P07643 homarus gammarus
30	57.8						Q93377 homarus gammarus
31	57.6						Q9ubx7 homarus gammarus
32	57.2						Q9h2r5 homarus gammarus
33	570.5						Q60259 homarus gammarus

ALIGNMENTS						
RESULT 1						
ID	TRY2_CANFA		STANDARD;		PRY;	247 AA.
AC	P06872;					
DT	01-JAN-1988	(Rel. 06, Created)				
DT	01-JAN-1988	(Rel. 06, Last sequence update)				
DT	15-JUN-2002	(Rel. 14, Last annotation update)				
DE	Trypsin, anionic precursor (EC 3.4.21.4).					
OS	Canis familiaris (Dog).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=86284628; PubMed=3841794;					
RA	Plinsky S.D.; Laforge K.S.; Scheie G.;					
RT	"Differential regulation of trypsinogen mRNA translation: full-length mRNA sequences encoding two oppositely charged trypsinogen isoenzymes in the dog pancreas".					
RT	In Mol. Cell. Biol. 5:2659-2676(1985).					
RL	-1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-1-Xaa, Lys-1-Xaa.					
CC	-1- SUBCELLULAR LOCATION: Extracellular.					
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC	EMBL; M11589; AAA30899.1; PIR; A26273; TRDG.					
DR	HSSP; P00763; 1DPO.					
DR	MEROPS; S01_258; -.					
DR	InterPro; IPR001314; Chymotrypsin.					
DR	InterPro; IPR001254; Ser_protease_Try.					
DR	Pfam; PF00089; trypsin; 1.					
DR	SMART; SM00020; tryp_SPC; 1.					
DR	PROSITE; PS00240; TRIPSIN_DOM; 1.					
DR	PROSITE; PS00134; TRYSIN_HIS; 1.					
DR	PROSITE; PS00135; TRYNSIN_SER; 1.					
KW	Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal; Multigene family.					
KW	FT SIGNAL	1				15
FT	PROPEP	16				23
FT	CHAIN	24				247
FT	ACT_SITE	63				63
FT	ACT_SITE	107				107
FT	ACT_SITE	200				200
FT	DISULFID	30				160
FT	DISULFID	48				64
FT	DISULFID	132				233
FT	DISULFID	206				206

Query Match 83.7%; Score 1111; DB 1; Length 246;
 Best Local Similarity 81.3%; Pred. No. 1..2e-90;
 Matches 200; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

Oy 1 MNPLLTIAFLGLAVATPTDDDKRIVGGTCEENSPVQVSILNAGYHFCGGSLLSDQYVVS 60
 Db 1 MRALLFLALVGAVAFPVDDDKRIVGGTCEENSPVQVSILNAGYHFCGGSLLSDQYVVS 60
 Oy REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 FT SITE 194 194
 FT CONFLICT 84 84 N -> D (IN REF. 1).
 FT SEQUENCE 246 AA; 26228 MW; A8D363080AAE666 CRG64;
 SQ ~~~~~~ 88 88 V -> I (IN REF. 1).

Query Match 83.7%; Score 1111; DB 1; Length 246;
 Best Local Similarity 81.3%; Pred. No. 1..2e-90;
 Matches 200; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

Oy 1 ARVATISLPRACAAAGTCOLISPGWNTLSSGTNPYPELQLCDAPITQAQCBASTYPOIT 180
 Db 121 ARVATVLPSSCAPAGTCOLISPGWNTLSSGTNPYPELQLCDAPITQAQCBASTYPOIT 180
 Oy 61 AAHCYKSRQIYRVLGEYNIDVLEGENQFINSAKVIRHNNYSNSWTLNDIMLKLSSPAVLN 120
 ~~~~~~ 61 AAHCYKSRQIYRVLGEYNIDVLEGENQFINSAKVIRHNNYSNSWTLNDIMLKLSSPAVLN 120  
 Oy 121 ARVATISLPRACAAAGTCOLISPGWNTLSSGTNPYPELQLCDAPITQAQCBASTYPOIT 180  
 Db 121 ARVATVLPSSCAPAGTCOLISPGWNTLSSGTNPYPELQLCDAPITQAQCBASTYPOIT 180  
 Oy 181 ENMICAGFLEGKDSCGDGGPVPVYCNCELOGIVSNGYCAOKNKPGVYTKYCNFVWIQ 240  
 Db 181 DMWCVCFLEGKDSCGDGGPVPVYCNCELOGIVSNGYCAOKNKPGVYTKYCNFVWIQ 240  
 Oy 241 STIAAN 246  
 Db 241 DTIAAN 246

RESULT 5  
 ID TRY1\_RAT STANDARD PRT; 246 AA.  
 AC P00765;  
 DT 21-JUL-1986 (Ref. 01, Created)  
 DT 15-JUN-2002 (Ref. 41, Last sequence update)  
 DE Trypsin I, anionic precursor (EC 3.4.21.4) (Pretrypsinogen 1).  
 GN TRY1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId:10116;  
 RN [1] SEQUENCE FROM N.A.  
 RN STRAIN="Sprague-Dawley";  
 RX MEDLINE=8226562; PubMed=6896710;  
 RA McDonald R.J., Stary S.J., Swift G.H.;  
 RA Rutter W.J.;  
 FT "two similar, but nonallelic rat pancreatic trypsinogens. Nucleotide  
 sequences of the cloned cDNAs";  
 RL J. Biol. Chem. 259:14255-14264(1984).  
 RN [2] SEQUENCE FROM N.A.  
 RX MEDLINE=85054888; PubMed=6094547;  
 RA Craik C.S., Choo Q.L., Quinto C., McDonald R.J.,  
 RA Rutter W.J.;  
 FT Structure of two related rat pancreatic trypsin genes.";  
 RL J. Biol. Chem. 259:14255-14264(1984).  
 RN [3] X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=87292123; PubMed=3112942;  
 RA Sprang S.; Standing T., Fletterick R.J., Stroud R.M., Finer-Moore J.,  
 RA Xuong N.-H., Hamlin R., Rutter W.J., Craik C.S.;  
 FT "The three-dimensional structure of Asn102 mutant of trypsin: role of  
 Asp102 in serine protease catalysis.";  
 RR Science 237:905-909(1987);  
 RL "- CATALYTIC ACTIVITY: Preferential cleavage: Arg-l-Xaa, Lys-l-Xaa.  
 CC -I- SUBCELLULAR LOCATION: Extracellular.

-1- MISCELLANEOUS: THIS SEQUENCE REPRESENTS THE PRECURSOR OF THE MAJOR FORM OF TRYPSIN PRODUCED BY THE ADULT PANCREAS.  
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions. Usage is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 CC DR EMBL: V01273; CAA24580 1; -.  
 CC DR EMBL: J0078; AAA98518 1; -.  
 CC DR PIR: A00948; TRTR1.  
 CC DR PDB: 1PTM; 15-JUL-93.  
 CC DR PDB: 2TRM; 16-JUL-98.  
 CC DR PDB: 1BRA; 30-APR-94.  
 CC DR PDB: 1BRR; 31-JUL-94.  
 CC DR PDB: 1BRC; 31-MAY-94.  
 CC DR MEROPS: S01.094; -.  
 CC DR InterPro: IPI001314; Chymotrypsin.  
 CC DR InterPro: IPI001254; ser\_protease\_TRY.  
 CC DR Pfam: PF00089; trypsin\_1.  
 CC DR PRINTS: PRO0722; CHYMOTRYPSIN.  
 CC DR SMART: SM00020; TRYPSIN\_1.  
 CC DR PROSITE: PS00040; TRYPSIN\_DOM; 1.  
 CC DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 CC DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 CC KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;  
 CC KW Multigene family; 3D-structure.  
 CC FT SIGNAL 1 15 ACTIVATION PEPTIDE.  
 CC FT PROPEP 16 23 TRYPSIN\_I.  
 CC FT CHAIN 24 246 ANIONIC.  
 CC FT ACT\_SITE 63 63 CHARGE RELAY SYSTEM.  
 CC FT ACT\_SITE 107 107 CHARGE RELAY SYSTEM.  
 CC FT DISULFID 200 200 CHARGE RELAY SYSTEM.  
 CC FT DISULFID 30 160  
 CC FT DISULFID 48 64  
 CC FT DISULFID 132 233  
 CC FT DISULFID 139 206  
 CC FT DISULFID 171 185  
 CC FT DISULFID 196 194  
 CC FT SITE 194 194 REQUIRED FOR SPECIFICITY.  
 CC FT STRAND 25 25  
 CC FT STRAND 28 29  
 CC FT TURN 32 33  
 CC FT TURN 36 37  
 CC FT STRAND 38 42  
 CC FT STRAND 46 54  
 CC FT TURN 55 56  
 CC FT STRAND 57 60  
 CC FT HELIX 62 64  
 CC FT TURN 70 73  
 CC FT STRAND 77 77  
 CC FT TURN 78 79  
 CC FT STRAND 86 95  
 CC FT TURN 97 98  
 CC FT STRAND 100 100  
 CC FT TURN 101 104  
 CC FT STRAND 105 105  
 CC FT TURN 106 106  
 CC FT STRAND 109 113  
 CC FT STRAND 127 127  
 CC FT TURN 135 136  
 CC FT STRAND 138 143  
 CC FT STRAND 157 157  
 CC FT STRAND 159 165  
 CC FT HELIX 168 174  
 CC FT TURN 176 178  
 CC FT STRAND 181 182  
 CC FT STRAND 183 186  
 CC FT TURN 189 190

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                    |                        |                                                                                                                                                                                                                     |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|--------------------|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| FT  | STRAND                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 194                                                               | 194                |                        | DR MEROPS; S01_056;                                                                                                                                                                                                 |
| FT  | TURN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 197                                                               | 198                |                        | DR InterPro; IPR001314; Chymotrypsin.                                                                                                                                                                               |
| FT  | TURN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 200                                                               | 201                |                        | DR InterPro; IPR001254; Ser_protease_Try.                                                                                                                                                                           |
| FT  | STRAND                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 203                                                               | 206                |                        | DR Pfam; PF00089; trypsin; 1.                                                                                                                                                                                       |
| FT  | TURN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 207                                                               | 208                |                        | DR PRINTS; PRO0722; CHYMTTRYPSIN.                                                                                                                                                                                   |
| FT  | STRAND                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 209                                                               | 216                |                        | DR SMART; SM00020; TRYSPC; 1.                                                                                                                                                                                       |
| FT  | TURN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 223                                                               | 224                |                        | DR PROSITE; PSS0240; TRYPSIN_DOM; 1.                                                                                                                                                                                |
| FT  | STRAND                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 227                                                               | 231                |                        | DR PROSITE; PS00134; TRYPSIN_HIS; 1.                                                                                                                                                                                |
| FT  | HELIX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 232                                                               | 234                |                        | DR PROSITE; PS00135; TRYPSIN_SER; 1.                                                                                                                                                                                |
| FT  | HELIX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 236                                                               | 244                |                        | KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;                                                                                                                                                |
| FT  | TURN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 245                                                               | 245                |                        | KW Multicene family.                                                                                                                                                                                                |
| SQ  | SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 246 AA:                                                           | 2595 MW:           | 6AFA0AD11943FB5 CRC64; | FT SIGNAL 1 15 ACTIVATION PEPTIDE.                                                                                                                                                                                  |
|     | Query Match                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 83.0%                                                             | Score 1101;        | DB 1; Length 246;      | FT PROPEP 16 24 TRYP SIN III, CATIONIC.                                                                                                                                                                             |
|     | Best Local Similarity                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 80.5%                                                             | Pred. No. 8.9e-90; |                        | FT ACT_CHAIN 25 247 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                            |
|     | Matches                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 198;                                                              | Mismatches         | 26; Indels 0; Gaps 0;  | FT ACT_SITE 64 64 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                              |
| QY  | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | MNPLLIAFLGAAVATPDDDKIVGGTYCEENSYPPQVSLSNAGYHFCGGSSLSDPQWVVS 60    |                    |                        | FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                            |
| QY  | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | :      :       :       :       :       :       :       :       :  |                    |                        | FT DISULFID 201 201 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                            |
| -Db | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | MSALLILALVGAAYAAPLEDDDKIVGGTCAPEHSPVYQVSLSNSGHFCGGSLNDQWVVS 60    |                    |                        | FT DISULFID 31 161 BY SIMILARITY.                                                                                                                                                                                   |
| '   | 61                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | AAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTNSWILDIMIKLSSPAVLN 120     |                    |                        | FT DISULFID 49 65 BY SIMILARITY.                                                                                                                                                                                    |
| Db  | 61                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | AAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTSSWTLNNDIMIKLSSPAVLN 120   |                    |                        | FT DISULFID 133 234 BY SIMILARITY.                                                                                                                                                                                  |
| QY  | 121                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | ARVATISLPRACAAPTQCLISGNLTSGGNTLQQDAPITLQDCEASYSPQQT 180           |                    |                        | FT DISULFID 140 207 BY SIMILARITY.                                                                                                                                                                                  |
| Db  | 121                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | ARVAPVALSACAAGTQCLISGNLTSGGNTLQQDAPITLQDCEASYSPQQT 180            |                    |                        | FT DISULFID 172 186 BY SIMILARITY.                                                                                                                                                                                  |
| QY  | 181                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | ENMICAGFLEGKGDSCQGDGGPQYVCNGELOGIYSGWYGCQAQKNKPQGYTAKVCFNVDWI 240 |                    |                        | FT DISULFID 197 221 BY SIMILARITY.                                                                                                                                                                                  |
| Db  | 181                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | SSMCVGFLEGKGDSCQGDGGPQYVCNGQLOGIVSGWYGCALPDNGTIVCNFGWQI 240       |                    |                        | FT SITE 195 195 REQUIRED FOR SPECIFICITY (BY SIMILARITY).                                                                                                                                                           |
| QY  | 241                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | STIAN 246                                                         |                    |                        | SQ SEQUENCE 247 AA: 26269 MW: D74892BAAS84E4AB CRC44;                                                                                                                                                               |
| Db  | 241                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | DTIAAN 246                                                        |                    |                        |                                                                                                                                                                                                                     |
|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                    |                        | Query Match 82.8%; Score 1098.5; DB 1; Length 247;                                                                                                                                                                  |
|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                    |                        | Best Local Similarity 78.5%; Pred. No. 1.5e-89; Matches 194; Conservative 31; Mismatches 21; Indels 1; Gaps 1;                                                                                                      |
|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                    |                        | Matches 194; Conservatve 31; Mismatches 21; Indels 1; Gaps 1;                                                                                                                                                       |
| QY  | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | MNPLLIAFLGAAVATPDDDKIVGGTYCEENSYPPQVSLSNAGYHFCGGSSLSDPQWVVS 59    |                    |                        | Qy 1 MNPLLIAFLGAAVATPDDDKIVGGTYCEENSYPPQVSLSNAGYHFCGGSSLSDPQWVVS 59                                                                                                                                                 |
| Db  | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | MAKALIFLAFQVALPLDDDKIVGGTYCQVSLSNAGYHFCGGSSLSDPQWVVS 60           |                    |                        | Qy 1 MKALIFLAFQVALPLDDDKIVGGTYCQVSLSNAGYHFCGGSSLSDPQWVVS 60                                                                                                                                                         |
| QY  | 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | SAAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTNSWILDIMIKLSSPAVLN 119    |                    |                        | Db 1 SAAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTNSWILDIMIKLSSPAVLN 119                                                                                                                                                 |
| Db  | 61                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | SAAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTSSWTLNNDIMIKLSSPAVLN 120  |                    |                        | Qy 60 SAAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTSSWTLNNDIMIKLSSPAVLN 120                                                                                                                                              |
| QY  | 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | NARVATISLPRACAAPTQCLISGNLTSGGNTLQQDAPITLQDCEASYSPQQT 179          |                    |                        | Db 61 SAAHCYKSRQIQRGLGYNNIDLEGNEQFINSAKVIKHPNTSSWTLNNDIMIKLSSPAVLN 120                                                                                                                                              |
| Db  | 121                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | :      :       :       :       :       :       :       :          |                    |                        | Qy 120 NARVATISLPRACAAPTQCLISGNLTSGGNTLQQDAPITLQDCEASYSPQQT 179                                                                                                                                                     |
| AC  | TRY3_RAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | STANDARD;                                                         | PRT;               | 247 AA.                | Db 121 NSRVSTYSLPRSCGSSTKCLVSGWNTLTSSTNPYPSLQCLAPVLSSCSSSYPSKI 180                                                                                                                                                  |
| AC  | P08426;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| DT  | 01-AUG-1988                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | (Rel. 08; Created)                                                |                    |                        | AC P07478; ID TRY2_HUMAN STANDARD; PRT; 247 AA.                                                                                                                                                                     |
| DT  | 01-AUG-1988                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | (Rel. 08; Last sequence update)                                   |                    |                        | DT 01-APR-1988 (Rel. 07, Created)                                                                                                                                                                                   |
| DT  | 15-JUN-2002                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | (Rel. 41; Last annotation update)                                 |                    |                        | DT 01-APR-1988 (Rel. 07, Last sequence update)                                                                                                                                                                      |
| DE  | TRYPSIN III, cationic precursor (EC 3.4.21.4)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | (Pretrypsinogen III).                                             |                    |                        | DT 15-JUN-2002 (Rel. 41, Last annotation update)                                                                                                                                                                    |
| GN  | TRY3.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                   |                    |                        | DE TRYPSIN II precursor (EC 3.4.21.4) (Anionic trypsinogen).                                                                                                                                                        |
| OS  | Rattus norvegicus (Rat).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                   |                    |                        | GN PRSS2 OR TRY2 QR TRYX2.                                                                                                                                                                                          |
|     | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus; NCBI-TaxID=10116;                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                   |                    |                        | OS Homo sapiens (Human).                                                                                                                                                                                            |
| JX  | [1]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                   |                    |                        | OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606;                                                                    |
| RN  | SEQUENCE FROM N.A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                   |                    |                        | RN [1]                                                                                                                                                                                                              |
| RX  | MEDLINE-8721609; PubMed=3607011;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                   |                    |                        | RP SEQUENCE FROM N.A.                                                                                                                                                                                               |
| RX  | Fletcher T.S., Ahadoff M., Craik C.S., Largman C.; "Isolation and characterization of a cDNA encoding rat cationic trypsinogen." Biochemistry 26:3081-3086 (1987).                                                                                                                                                                                                                                                                                                                                                                      |                                                                   |                    |                        | RX MEDLINE=8622112; PubMed=3011602;                                                                                                                                                                                 |
| RR  | - - CARBILYTIC ACTIVITY: Preferential cleavage: Arg- -xaa, Lys- -xaa.                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                   |                    |                        | RA Eimi M., Nakamura Y., Ogawa M., Yamamoto T., Nishide T., Mori T., Matsubara K.; "Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens." Gene 41:305-310(1986). |
| CC  | - - SUBCELLULAR LOCATION: Extracellular.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                   |                    |                        | RR [2]                                                                                                                                                                                                              |
| CC  | - - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| CC  | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| CC  | CC This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).               |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| CC  | DR M16624; AAA41985.1; -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| DR  | PIR: A27547; A27547.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| DR  | HSSP; P00763; 1DPO.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                   |                    |                        |                                                                                                                                                                                                                     |
| RN  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                    |                        |                                                                                                                                                                                                                     |

|                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                     |
|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RP                        | SEQUENCE OF 16-49.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PRT; 247 AA.                                                                                                                                                                                        |
| RA                        | Medline=9091010; PubMed=2598466;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | AC P07477; Q92955; Q9HAN4; Q9HAN6; Q9HAN7;                                                                                                                                                          |
| RA                        | Kimland M., Russick C., Maris W.H., Borgstrom A.;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | AC 01-Apr-1988 (Rel. 07, Created)                                                                                                                                                                   |
| RA                        | "Immunoactive anionic and cationic trypsin in human serum.";                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | DT 01-Apr-1988 (Rel. 07, Last sequence update)                                                                                                                                                      |
| RL                        | Clin. Chim. Acta 184:31-46 (1989).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | DT 01-JUN-2002 (Rel. 41, Last annotation update)                                                                                                                                                    |
| CC                        | -1 - CATALYTIC ACTIVITY: Preferential cleavage: Arg-1-Xaa, Lys-1-Xaa.                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | DE Trypsin I precursor (EC 3.4.21.4) (Cationic trypsinogen).                                                                                                                                        |
| CC                        | -1 - SUBCELLULAR LOCATION: Extracellular.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | GN PRSS1 OR TRY1 OR TRPL OR TRY1.                                                                                                                                                                   |
| CC                        | -1 - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | OS Homo sapiens (Human).                                                                                                                                                                            |
| CC                        | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to licensee@isb-sib.ch). | OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                                                                     |
| DR                        | EMBL: M27602; AAA61232.1; HSSP: P00763; B25852.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | OC TAXID=9606;                                                                                                                                                                                      |
| DR                        | MEROPS: S01_258; -;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | OC [1]                                                                                                                                                                                              |
| DR                        | GeneID: HGNC:9483; PRSS2.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | RP SEQUENCE FROM N.A.                                                                                                                                                                               |
| DR                        | MTM; 601564;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | RX MEDLINE=90091010; PubMed=2598466;                                                                                                                                                                |
| DR                        | InterPro; IPR001314; Chymotrypsin.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | RA Kimland M., Russick C., Marks W.H., Borgstrom A.;                                                                                                                                                |
| DR                        | InterPro; IPR001254; Ser_protease_Try.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | RT "Immunoreactive anionic and cationic trypsin in human serum.";                                                                                                                                   |
| DR                        | PFam; PF00890; Trypsin; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | RL Clin. Chim. Acta 184:31-46(1989).                                                                                                                                                                |
| DR                        | PRINTS: PRO0722; CHYMOTRYPSIN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | RN [3]                                                                                                                                                                                              |
| DR                        | SMART: SM00020; TRYSPC; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | RP SEQUENCE OF 68-151 FROM N.A. AND VARIANT HPC HIS-122.                                                                                                                                            |
| DR                        | PS50240; TRYSPIN_DOM; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | RX MEDLINE=96438847; PubMed=8841182;                                                                                                                                                                |
| DR                        | PROSITE; PS00134; TRYSPIN_HIS; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | RA Whitcomb D.C., Gorry M.C., Preston R.A., Furey W., Sossenheimer M.J., Ulrich C.D., Martin S.P., Gates L.K. Jr., Anann S.T., Toskes P.P., Liddle R., McGrath K., Domo G., Post J.C., Enrich G.D.; |
| DR                        | PROSITE; PS00135; TRYSPIN_SER; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | RT "Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene."                                                                                                              |
| KW                        | Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal; Multigene family.                                                                                                                                                                                                                                                                                                                                                                                                                                                      | RT Nat. Genet. 14:141-145 (1996).                                                                                                                                                                   |
| FT                        | SIGNAL 1 15                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.                                                                                                                                          |
| FT                        | PROPEP 16 23                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | RN [4]                                                                                                                                                                                              |
| FT                        | CHAIN 24 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | RP SEQUENCE OF 15-67 FROM N.A. AND VARIANT HPC GLY-22.                                                                                                                                              |
| FT                        | ACT_SITE 63 63                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | RX MEDLINE=20389962; PubMed=1093331;                                                                                                                                                                |
| FT                        | ACT_SITE 107 107                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | RA Teich N., Ockeniga J., Hoffmeyer A., Manns M., Mossner J., Keim V.;                                                                                                                              |
| FT                        | ACT_SITE 200 200                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | RT "Chronic pancreatitis associated with an activation peptide mutation that facilitates trypsin activation.";                                                                                      |
| FT                        | DISULFID 30 160                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | RT RL Gastroenterology 119:461-465 (2000).                                                                                                                                                          |
| FT                        | DISULFID 48 64                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | RN [5]                                                                                                                                                                                              |
| FT                        | DISULFID 171 185                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | RP SEQUENCE OF 68-151 FROM N.A. AND VARIANT HPC P-104; C-116 AND F-139.                                                                                                                             |
| FT                        | DISULFID 196 220                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | RA Teich N., Bauer N., Mossner J., Keim V.;                                                                                                                                                         |
| FT                        | SITE 194 194                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.                                                                                                                                          |
| SQ                        | SEQUENCE 247 AA; 26488 NW; 82B0F1EB8ED5DB CRC64;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | RN [6]                                                                                                                                                                                              |
| Query Match               | Score 1097; DB 1; Length 247;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), AND PHOSPHORYLATION.                                                                                                                                      |
| Best Local Similarity     | 81.8%; Pred. No. 2e-89;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | RX MEDLINE=9626696; PubMed=86683601;                                                                                                                                                                |
| Matches 202; Conservative | 21; Mismatches 24; Indels 0; Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | RA Gaboriaud C., Serre L., Guy-Crotte O., Forest E., Fontecilla-Camps J.-C.;                                                                                                                        |
| Oy                        | 1 MNPLLILAFLLGAAVATPDDDKIVGGTCEENSVPYOVSNLNAGYHFCGGSLISDOWVYS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | RT "Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr151.";                                                                                                                   |
| Db                        | 1 MNPLLILTFVAAVAAPPDDDKIVGGTCEENSVPYOVSNLNAGYHFCGGSLISEQWVVS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | RT J. Mol. Biol. 259:995-1010 (1996).                                                                                                                                                               |
| Qy                        | 61 AAHCYKSR1QVRGEYNVDLGEQEINSAKVIRHPNNWSILDNDIMLKSSLPAVLN 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | RP VARIANTS HPC ILE-29 AND HIS-122.                                                                                                                                                                 |
| Db                        | 61 AGHCYKSR1QVRGEYNVDLGEQEINSAKVIRHPNNWSILDNDIMLKSSLPAVLN 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | RX MEDLINE=98463197; PubMed=9322498;                                                                                                                                                                |
| Qy                        | 61 ARVATISPRACAPGTCOLISGMNTLSSGTNPVLLQCLDAPIITQACEASPGQT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | RA Goryt M.C., Gabbaizadeh D., Furey W., Gates L.K. Jr., Preston R.A., Aston C.E., Zhang Y., Ulrich C., Ehrlich G.D., Whitcomb D.C.;                                                                |
| Db                        | 121 SRVSAISLPTAPPAGTSESLISGWNTLSSGADYPDEQLCLDAPIVSLRECEASPGKT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | RT "Mutations of the cationic trypsinogen gene are associated with recurrent acute and chronic pancreatitis.";                                                                                      |
| Qy                        | 181 ENMICAGPLEGGKRDSCQGDGGPPVYCNGELOGYIWSMGCAOKNPKGVYTKVCFNFDWQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | RL Gastroenterology 113:1063-1068 (1997).                                                                                                                                                           |
| Db                        | 181 NNRFVCVGPLEGGKRDSCQGDGGPPVYSNGELQGIVSWGTVQKRNPGVYKVNWDNIK 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | RN [8]                                                                                                                                                                                              |
| Qy                        | 241 SIAAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | RP VARIANT HPC VAL-16 AND HIS-122.                                                                                                                                                                  |
| Db                        | 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | RX MEDLINE=99155444; PubMed=10381903;                                                                                                                                                               |
| RESULT 8                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | RA Witt H., Luck W., Becker M.;                                                                                                                                                                     |
| TRY1_HUMAN                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | RT "A signal peptide cleavage site mutation in the cationic trypsinogen gene is strongly associated with chronic pancreatitis.";                                                                    |



| REQUIRED FOR SPECIFICITY (BY SIMILARITY). |                                                                                                                                                                                                                      |
|-------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RL                                        | Biochim. Biophys. Acta 69:115-129(1963).                                                                                                                                                                             |
| RN [2]                                    | SEQUENCE OF 9-231.                                                                                                                                                                                                   |
| RN                                        | MEDLINE=72258692; PubMed=738933;                                                                                                                                                                                     |
| RX                                        | "Determination of the amino acid sequence of porcine trypsin by sequenator analysis.";                                                                                                                               |
| RA                                        | Hermanson M.A., Ericsson L.H., Neurath H., Walsh K.A.;                                                                                                                                                               |
| RT                                        | "Refinement of the amino acid sequence of porcine trypsin by sequenator analysis.";                                                                                                                                  |
| RT                                        | Biochemistry 12:3146-3153(1973).                                                                                                                                                                                     |
| RN [3]                                    | X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).                                                                                                                                                                               |
| RX                                        | MEDLINE=93187598; PubMed=8445634;                                                                                                                                                                                    |
| RA                                        | Huang Q., Liu S., Tang Y.;                                                                                                                                                                                           |
| RT                                        | "Refined 1.6-A resolution crystal structure of the complex formed between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the squash family. Detailed comparison with bovine beta-trypsin and its complex.;" |
| RT                                        | J. Mol. Biol. 229:1022-1030(1993).                                                                                                                                                                                   |
| RN [4]                                    | X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).                                                                                                                                                                               |
| RP                                        | MEDLINE=22201369; PubMed=5151419;                                                                                                                                                                                    |
| RA                                        | Huang Q., Liu S., Tang Y., Zeng F., Qian R.;                                                                                                                                                                         |
| RT                                        | "Amino acid sequencing of a trypsin inhibitor by refined 1.6 A x-ray crystal structure of its complex with porcine beta-trypsin.;"                                                                                   |
| RT                                        | FEBS Lett. 297:143-146(1992).                                                                                                                                                                                        |
| RN [5]                                    | X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).                                                                                                                                                                               |
| RP                                        | MEDLINE=95035057; PubMed=7949885;                                                                                                                                                                                    |
| RA                                        | Huang Q., Wang Z., Li Y., Liu S., Tang Y.;                                                                                                                                                                           |
| RT                                        | "Refined 1.8-A resolution crystal structure of the porcine epsilon-trypsin.;"                                                                                                                                        |
| RT                                        | FEBS Lett. 297:143-146(1992).                                                                                                                                                                                        |
| RN [6]                                    | X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH LDTI.                                                                                                                                                          |
| RP                                        | MEDLINE=97390427; PubMed=9242660;                                                                                                                                                                                    |
| RA                                        | Stubbs M.T., Morenweiser R., Sturzebecher J., Bauer M., Bode W., Huber R., Piechotta G.P., Sommerhoff C.P., Fritz H., Auerswald E.A.;                                                                                |
| RT                                        | "The three-dimensional structure of recombinant leech-derived trypsinase inhibitor in complex with trypsin. Implications for the structure of human mast cell trypsinase and its inhibition.";                       |
| RT                                        | J. Biol. Chem. 272:19931-19937(1997).                                                                                                                                                                                |
| RN [7]                                    | X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) OF COMPLEX WITH LDTI.                                                                                                                                                         |
| RP                                        | MEDLINE=98046092; PubMed=9384562;                                                                                                                                                                                    |
| RA                                        | di Marco S., Priestle J.P.;                                                                                                                                                                                          |
| RT                                        | "Structure of the complex of leech-derived trypsinase inhibitor (LDTI) with trypsin and modelling of the LDTI-trypsin system.";                                                                                      |
| RT                                        | Structure 5:1465-1474(1997).                                                                                                                                                                                         |
| --                                        | "-- CATALYTIC ACTIVITY: Preferential cleavage: Arg- -xaa, Lys- -xaa.                                                                                                                                                 |
| --                                        | "-- SIMILARITY: Extracellular location; Extracellular.                                                                                                                                                               |
| CC                                        | PIR: A00947; TRPTR.                                                                                                                                                                                                  |
| DR                                        | PDB: 1MCT; 31-JAN-94.                                                                                                                                                                                                |
| DR                                        | PDB: 1AKS; 12-FEB-97.                                                                                                                                                                                                |
| DR                                        | PDB: 1EPT; 07-FEB-95.                                                                                                                                                                                                |
| DR                                        | PDB: 1TPX; 21-JAN-98.                                                                                                                                                                                                |
| DR                                        | PDB: 1LDT; 20-MAY-98.                                                                                                                                                                                                |
| DR                                        | PDB: 1AND; 01-JUL-98.                                                                                                                                                                                                |
| DR                                        | PDB: 1AWW; 18-NOV-98.                                                                                                                                                                                                |
| DR                                        | PDB: 1AYW; 18-NOV-98.                                                                                                                                                                                                |
| DR                                        | MEROPS: S01_151; ";                                                                                                                                                                                                  |
| DR                                        | InterPro: IPR001314; Chymotrypsin.                                                                                                                                                                                   |
| DR                                        | InterPro: IPR01254; Ser-protease_trypsin.                                                                                                                                                                            |
| DR                                        | Pfam: PF00089; trypsin_1.                                                                                                                                                                                            |
| PRINTS: PR00732; CHYMOTRYPSIN.            |                                                                                                                                                                                                                      |
| SMART: SM0020; tryp_sp; 1.                |                                                                                                                                                                                                                      |
| PROSITE: PS50240; TRYPSIN_DOM; 1.         |                                                                                                                                                                                                                      |
| PROSITE: PS00134; TRYPSIN_HIS; 1.         |                                                                                                                                                                                                                      |
| PROSITE: PS00135; TRYPSIN_SER; 1.         |                                                                                                                                                                                                                      |
| KW                                        | Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; 3D-structure.                                                                                                                                              |
| ACT_PEP                                   | 1 8                                                                                                                                                                                                                  |
| CHAIN                                     | 9 231                                                                                                                                                                                                                |
| FT                                        | ACT_SITE 48 48                                                                                                                                                                                                       |
| FT                                        | ACT_SITE 92 92                                                                                                                                                                                                       |
| FT                                        | ACT_SITE 185 185                                                                                                                                                                                                     |
| FT                                        | DISULFID 15 145                                                                                                                                                                                                      |
| FT                                        | DISULFID 33 49                                                                                                                                                                                                       |
| FT                                        | DISULFID 117 218                                                                                                                                                                                                     |
| FT                                        | DISULFID 124 191                                                                                                                                                                                                     |
| FT                                        | DISULFID 156 170                                                                                                                                                                                                     |
| FT                                        | DISULFID 181 205                                                                                                                                                                                                     |
| FT                                        | SITE 179 179                                                                                                                                                                                                         |
| FT                                        | VARIANT 20 20                                                                                                                                                                                                        |
| FT                                        | STRAND 10 10                                                                                                                                                                                                         |
| FT                                        | STRAND 13 14                                                                                                                                                                                                         |
| FT                                        | TURN 17 18                                                                                                                                                                                                           |
| FT                                        | TURN 21 22                                                                                                                                                                                                           |
| FT                                        | STRAND 23 27                                                                                                                                                                                                         |
| FT                                        | STRAND 31 39                                                                                                                                                                                                         |
| FT                                        | TURN 40 41                                                                                                                                                                                                           |
| FT                                        | STRAND 42 45                                                                                                                                                                                                         |
| FT                                        | HELIX 47 49                                                                                                                                                                                                          |
| FT                                        | STRAND 55 58                                                                                                                                                                                                         |
| FT                                        | STRAND 62 62                                                                                                                                                                                                         |
| FT                                        | TURN 63 64                                                                                                                                                                                                           |
| FT                                        | STRAND 71 80                                                                                                                                                                                                         |
| FT                                        | TURN 82 83                                                                                                                                                                                                           |
| FT                                        | TURN 86 88                                                                                                                                                                                                           |
| FT                                        | TURN 90 91                                                                                                                                                                                                           |
| FT                                        | STRAND 94 98                                                                                                                                                                                                         |
| FT                                        | STRAND 112 112                                                                                                                                                                                                       |
| FT                                        | TURN 120 121                                                                                                                                                                                                         |
| FT                                        | STRAND 123 128                                                                                                                                                                                                       |
| FT                                        | STRAND 142 142                                                                                                                                                                                                       |
| FT                                        | STRAND 144 150                                                                                                                                                                                                       |
| FT                                        | HELIX 153 159                                                                                                                                                                                                        |
| FT                                        | TURN 161 163                                                                                                                                                                                                         |
| FT                                        | TURN 166 167                                                                                                                                                                                                         |
| FT                                        | STRAND 168 171                                                                                                                                                                                                       |
| FT                                        | TURN 174 175                                                                                                                                                                                                         |
| FT                                        | STRAND 179 179                                                                                                                                                                                                       |
| FT                                        | TURN 182 183                                                                                                                                                                                                         |
| FT                                        | TURN 185 186                                                                                                                                                                                                         |
| FT                                        | STRAND 188 191                                                                                                                                                                                                       |
| FT                                        | TURN 192 193                                                                                                                                                                                                         |
| FT                                        | STRAND 194 202                                                                                                                                                                                                       |
| FT                                        | STRAND 207 207                                                                                                                                                                                                       |
| FT                                        | TURN 208 209                                                                                                                                                                                                         |
| FT                                        | STRAND 210 210                                                                                                                                                                                                       |
| FT                                        | STRAND 212 216                                                                                                                                                                                                       |
| FT                                        | HELIX 217 219                                                                                                                                                                                                        |
| FT                                        | HELIX 221 230                                                                                                                                                                                                        |
| SO                                        | SEQUENCE 231 AA; 24409 MW; A0A125CF7/BC138C2·CRRC64;                                                                                                                                                                 |
| Query Match                               | 80.7%                                                                                                                                                                                                                |
| Best Local Similarity                     | 83.0%                                                                                                                                                                                                                |
| Matches                                   | 191; Conservative 21; Mismatches 18; Indels 0; Gaps                                                                                                                                                                  |
| Qy                                        | 17 PDDDDKIVGGWNTLSSGNNYPPELQLDAPILQAQCERSYPGQITENMICAGLEGGKDSC 1                                                                                                                                                     |
| Db                                        | 2 PDDDDKIVGGTCAANSIPIGVSLNSGSHFCGGSLINSQWVTSAAACYSKRIQVRGEH 6                                                                                                                                                        |
| Qy                                        | 77 NDVLEGEQNEQFINSAKVIRHPDYNWSWILNDNDMLIKLSSPAVLNARYATISLPRACAGP 1                                                                                                                                                   |
| Db                                        | 62 NDVLEGEQNEQFINSAKLTHPPENGTLNDNDMLIKLSSPAVLNSRVATSPRSCAAAG 1                                                                                                                                                       |
| Qy                                        | 137 TQLISGWGNLSSGNNYPPELQLDAPILQAQCERSYPGQITENMICAGLEGGKDSC 1                                                                                                                                                        |
| Db                                        | 122 TQLISGWGNLSSGNNYPPELQLDAPILQAQCERSYPGQITENMICAGLEGGKDSC 1                                                                                                                                                        |
| Qy                                        | 197 QDSGGPVCNGELOQIVSNGVGCDAOKRNPKGVYTKVCFNVDIQTOSTIAAN 246                                                                                                                                                          |
| Db                                        | 182 QDSGGPVCNGQLOQIVSNGVGCDAOKRNPKGVYTKVCFNVDIQTOSTIAAN 231                                                                                                                                                          |

| RESULT 10                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| TRY3_HUMAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | STANDARD; PRT; 247 AA. |
| AC P11951;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                        |
| DT 01-APR-1990 (Rel. 14, Created)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                        |
| DT 01-APR-1990 (Rel. 14, Last sequence update)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| DT 15-JUN-2002 (Rel. 41, Last annotation update)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| DE TRYPSIN III precursor (EC 3.4.21.4).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                        |
| GN PRSS3 OR TRY3.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                        |
| OS Homo sapiens (Human).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                        |
| OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| NCBI_TaxID=9606;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| RN [1]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| RP SEQUENCE FROM N.A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| RC TISSUE=Pancreas;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| RA MEDLINE=9022185; PubMed=2326201;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| RA Tani T., Kawashima I., Mita K., Takiguchi Y.;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| RA "Nucleotide sequence of human pancreatic trypsinogen III cDNA.";                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| RA Nucleic Acids Res. 18:161-161(1990).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                        |
| CA -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg -l-Xaa, Lys -l-Xaa.                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| CA -!- SUBCELLULAR LOCATION: Extracellular family S1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch)    |                        |
| DR EMBL; X115505; CAA35327.1; -.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| DR PIR; S12764; S12764.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                        |
| DR MEROPS; S01.174; -.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| DR HSSP; P00761; 1EPT.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| DR Genew; HGNC:9486; PRSS3.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                        |
| DR InterPro; IPRO01314; Chymotrypsin.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| DR InterPro; IPRO01254; Ser_Protease_Try.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| DR Pfam; PF00089; trypsin; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| DR PRINTS; PR00722; CHYMOTRYPTIN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                        |
| DR SMART; SM00020; TRY_SP_C_1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| DR PROSITE; PS00240; TRYSPIN_DOM_1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| DR PROSITE; PS00134; TRYPSIN_HIS_1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| DR PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                        |
| KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal; Multigene family.                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| FT SIGNAL 1 15 ACTIVATION PEPTIDE.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| FT PROPEP 16 23 TRYPSIN III.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                        |
| FT ACT_CHAIN 24 247 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                        |
| FT ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| FT ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                        |
| FT ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                        |
| FT DISULFID 30 160 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                        |
| FT DISULFID 48 64 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| FT DISULFID 139 206 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| FT DISULFID 171 185 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| FT SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| SQ SEQUENCE 247 AA; 26776 MW; 697DE163FICEB06 CRC64;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                        |
| Query Match 79.7% Score 1058; DB 1; Length 247;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| Best Local Similarity 78.5%; Pred. No. 5.4e-86;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| Matches 194; Conservative 27; Mismatches 26; Indels 0; Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| QY 1 MNPLIILAFAGAATPTDDKIVGGTCCEPYOVSILNAGTHFCGSSLISDOWVYS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| 1 MNPLIILAFAGAATPTDDKIVGGTCCEPYOVSILNAGTHFCGSSLISDOWVYS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                        |
| 61 AHCKYKSRIQYRLGEYNIDVLEGNEQFINSAKVIRHPNNTSWILDNDIMLKSPAVLN 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| 61 AHCKYKSRIQYRLGEYNIDVLEGNEQFINSAKVIRHPNNTSWILDNDIMLKSPAVLN 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPTAPPAGTECLISGNTLSEFGADYDPDELKCLDAVLRENECKASCSPGKIT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| RESULT 11                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| TRY4_HUMAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | STANDARD; PRT; 304 AA. |
| ID TRY4_HUMAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| ID Q15665; Q9UQV7.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| AC P35030;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                        |
| DT 01-FEB-1994 (Rel. 28, Created)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                        |
| DT 01-FEB-1994 (Rel. 28, Last sequence update)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| DT 15-JUN-2002 (Rel. 41, Last annotation update)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| DE Trypsin IV precursor (EC 3.4.21.4) (Brain trypsinogen)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| DE (Mesotrypsinogen).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| DE (Mesotrypsinogen).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| GN PRSS4.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| OS Homo sapiens (Human).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                        |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                        |
| OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| NCBI_TaxID=9606;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| RN [1]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| RP SEQUENCE FROM N.A. (ISOFORMS A AND B).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| RA MEDIline=94123994; PubMed=8294000;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| RA Wiegand U., Corbach S., Minn A., Kang J., Mueller-Hill B.;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| RA Cloning of the cDNA encoding human brain trypsinogen and characterization of its product.";                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| RA Gene 136:167-175(1993).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                        |
| RN [2]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| RP SEQUENCE FROM N.A. (ISOFORM C).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| RA Fukuoka S.;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| RA Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                        |
| CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg -l-Xaa, Lys -l-Xaa.                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| CC -!- SUBCELLULAR LOCATION: Extracellular.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                        |
| CC -!- ALTERNATIVE PRODUCTS: 3 isoforms; A (shown here), B and C; seem to be produced by alternative splicing.                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| CC -!- TISSUE SPECIFICITY: PANCREAS AND BRAIN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). |                        |
| CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). |                        |
| CC EMBL; X72781; CB58178.1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                        |
| DR EMBL; X71345; CA50484.1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                        |
| DR EMBL; D45417; BRA08257.1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                        |
| DR PIR; S33496; S33496.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                        |
| DR HSSP; P0477; ITRN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| DR MEROPS; S01.174.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| DR InterPro; IPR00314; Chymotrypsin.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                        |
| DR Best; IPR001254; Ser_protease_Try.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| PFam; PF00089; trypsin; 1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                        |
| PRINTS; PR00722; CHYMOTRYPSIN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| SMART; SM00020; TRYSPIN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                        |
| PROSITE; PS50240; TRY_SP_C_1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| PROSITE; PS00134; TRYPSIN_HIS_1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| PROSITE; PS00240; TRYSPIN_DOM_1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal; Multigene family.                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| SIGNAL 1 15 ACTIVATION PEPTIDE.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| PROPEP 16 23 TRYPSIN III.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                        |
| ACT_CHAIN 24 247 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| ACT_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                        |
| ACT_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| ACT_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| DISULFID 30 160 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| DISULFID 48 64 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| DISULFID 139 206 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| DISULFID 171 185 BY SIMILARITY.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| SEQUENCE 247 AA; 26776 MW; 697DE163FICEB06 CRC64;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                        |
| Query Match 79.7% Score 1058; DB 1; Length 247;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| Best Local Similarity 78.5%; Pred. No. 5.4e-86;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                        |
| Matches 194; Conservative 27; Mismatches 26; Indels 0; Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                        |
| QY 1 MNPLIILAFAGAATPTDDKIVGGTCCEPYOVSILNAGTHFCGSSLISDOWVYS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                        |
| 1 MNPLIILAFAGAATPTDDKIVGGTCCEPYOVSILNAGTHFCGSSLISDOWVYS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                        |
| 61 AHCKYKSRIQYRLGEYNIDVLEGNEQFINSAKVIRHPNNTSWILDNDIMLKSPAVLN 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| 61 AHCKYKSRIQYRLGEYNIDVLEGNEQFINSAKVIRHPNNTSWILDNDIMLKSPAVLN 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 241 DTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                        |
| Db 121 ARVATISLPRCAAAGTQCLISIGNNTLSGGTNPYELLOCLDAPILTOACEASTYPOGT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                        |
| QY 181 ENMICAGFLEGGKDSQGDGGPVVCNGEQTIVSGYGAQKNGPQYATKVNFDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                        |
| Db 181 NSMFCVGFLLEGGKDSWKRDGGPVVCNGOLQGVTSWGHGCAWKNRPGVITKVNTVDWIQ 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                        |
| QY 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | </                     |

|                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
|----------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|
| PT                                                             | ACT_SITE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 164                                                               | 164                                                                     | CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                          | DR SMART; SM00020; TRYPSIN; SPC; 1.                        |
| PT                                                             | ACT_SITE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 257                                                               | 257                                                                     | CHARGE RELAY SYSTEM (BY SIMILARITY).                                                                                                                                                                                          | DR PROSITE; PS0240; TRYPSIN_DOM; 1.                        |
| PT                                                             | DISULFID                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 87                                                                | 217                                                                     | BY SIMILARITY.                                                                                                                                                                                                                | DR PROSITE; PS00134; TRYPSIN_HIS; 1.                       |
| PT                                                             | DISULFID                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 105                                                               | 121                                                                     | BY SIMILARITY.                                                                                                                                                                                                                | DR PROSITE; PS00135; TRYPSIN_SER; 1.                       |
| PT                                                             | DISULFID                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 196                                                               | 263                                                                     | BY SIMILARITY.                                                                                                                                                                                                                | KW Hydrolase; Serine protease; Digestion; Zymogen; Signal; |
| PT                                                             | DISULFID                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 228                                                               | 242                                                                     | BY SIMILARITY.                                                                                                                                                                                                                | KW Multigene family.                                       |
| PT                                                             | DISULFID                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 253                                                               | 277                                                                     | BY SIMILARITY.                                                                                                                                                                                                                | FT SIGNAL 1 15 BY SIMILARITY.                              |
| PT                                                             | DISULFID                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 251                                                               | 251                                                                     | REQUIRED FOR SPECIFICITY (BY SIMILARITY).                                                                                                                                                                                     | FT ACTIVATION PEPTIDE (BY SIMILARITY).                     |
| PT                                                             | VARSPLIC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 1                                                                 | 45                                                                      | MCGPDRCPAWPGPGRAVKCGKLARPARVERGAQRG<br>GAGL->M (IN ISOFORM B).                                                                                                                                                                | FT TRPSIN 22 244 BY SIMILARITY.                            |
| PT                                                             | VARSPLIC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 1                                                                 | 70                                                                      | MCGPDRCPAWPGPGRAVKCGKLAAARPGRVERGAQRG<br>GAGLEHPLLGRTRAARDADGCEALGT -> MNPFLLIA<br>FVGAA (IN ISOFORM C).                                                                                                                      | FT ACT_SITE 61 61 BY SIMILARITY.                           |
| PT                                                             | CONFLICT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 89                                                                | 89                                                                      | MISSING (IN REF. 1; CAA50484).                                                                                                                                                                                                | FT ACT_SITE 105 105 BY SIMILARITY.                         |
| SQ                                                             | SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 304 AA;                                                           | 32499 MW;                                                               | 4C4316C31F1D0FPC CRC64;                                                                                                                                                                                                       | FT DISULFID 28 158 BY SIMILARITY.                          |
| Query Match                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 78.1%                                                             | Score 1036;                                                             | DB 1; Length 244;                                                                                                                                                                                                             | FT DISULFID 46 62 BY SIMILARITY.                           |
| Best Local Similarity                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 80.3%                                                             | Pred. No. 5. 9e-84;                                                     | FT DISULFID 130 231 BY SIMILARITY.                                                                                                                                                                                            |                                                            |
| Matches 188; Conservative 24; Mismatches 22; Indels 0; Gaps 0; |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   | FT DISULFID 137 204 BY SIMILARITY.                                      |                                                                                                                                                                                                                               |                                                            |
| SQ                                                             | SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 244 AA;                                                           | 304 AA;                                                                 | 2079 MW;                                                                                                                                                                                                                      | FT DISULFID 169 183 BY SIMILARITY.                         |
| SQ                                                             | SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 244 AA;                                                           | 304 AA;                                                                 | 663F29CB3300B323 CRC64;                                                                                                                                                                                                       | FT SITE 194 218 BY SIMILARITY.                             |
| SQ                                                             | SEQUENCE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 244 AA;                                                           | 304 AA;                                                                 | 2079 MW;                                                                                                                                                                                                                      | FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).  |
| Query Match                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 77.8%                                                             | Score 1033;                                                             | DB 1; Length 244;                                                                                                                                                                                                             | FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).  |
| Best Local Similarity                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 77.2%                                                             | Pred. No. 8. 1e-84;                                                     | FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).                                                                                                                                                                     |                                                            |
| Matches 190; Conservative 26; Mismatches 28; Indels 2; Gaps 1; |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   | FT SITE 192 192 REQUIRED FOR SPECIFICITY (BY SIMILARITY).               |                                                                                                                                                                                                                               |                                                            |
| Qy                                                             | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | MNDLILAFGLAVAVATPDDDKRIVGGVTEENSVPYQVSNLNAGYHFCGSLISDQWVVS 60     | Qy 1 MNNDLILAFGLAVAVATPDDDKRIVGGVTEENSVPYQVSNLNAGYHFCGSLISDQWVVS 60     | Qy 1 MNNDLILAFGLAVAVATPDDDKRIVGGVTEENSVPYQVSNLNAGYHFCGSLISDQWVVS 60                                                                                                                                                           |                                                            |
| Db                                                             | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | MKFLVLVILGAAVA - FEDDDKLVGGFTCAKAVPVQYSNLNAGYHFCGSLISDQWVVS 58    | Db 1 MKFLVLVILGAAVA - FEDDDKLVGGFTCAKAVPVQYSNLNAGYHFCGSLISDQWVVS 58     | Db 1 MKFLVLVILGAAVA - FEDDDKLVGGFTCAKAVPVQYSNLNAGYHFCGSLISDQWVVS 58                                                                                                                                                           |                                                            |
| Qy                                                             | 61                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | AACYCKSRISQVRGEYNIDYLEGEQNEQFINSAKYTRHPNYSWILDNDIMLIKLSPPAVLN 120 | Qy 61 AACYCKSRISQVRGEYNIDYLEGEQNEQFINSAKYTRHPNYSWILDNDIMLIKLSPPAVLN 120 | Qy 61 AACYCKSRISQVRGEYNIDYLEGEQNEQFINSAKYTRHPNYSWILDNDIMLIKLSPPAVLN 120                                                                                                                                                       |                                                            |
| Db                                                             | 59                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | AACYCKSRISQVRGEHNITALNEGTEQFIDSQKVTKHPNNSRNLDNDIMLIKLSPPAVLN 118  | Db 59 AACYCKSRISQVRGEHNITALNEGTEQFIDSQKVTKHPNNSRNLDNDIMLIKLSPPAVLN 118  | Db 59 AACYCKSRISQVRGEHNITALNEGTEQFIDSQKVTKHPNNSRNLDNDIMLIKLSPPAVLN 118                                                                                                                                                        |                                                            |
| Qy                                                             | 121                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | ARYATISLPRACAAAPGTOCLISWGNTLSSGTNYPPELQLDAPLTQOCEASYPGQT 180      | Qy 121 ARYATISLPRACAAAPGTOCLISWGNTLSSGTNYPPELQLDAPLTQOCEASYPGQT 180     | Qy 121 ARYATISLPRACAAAPGTOCLISWGNTLSSGTNYPPELQLDAPLTQOCEASYPGQT 180                                                                                                                                                           |                                                            |
| Db                                                             | 119                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | ANIQSVLPSPSACASAGTNCNLISWGNTLSSGTNYPDLQLCNPALTDQSNSYPGEIT 178     | Db 119 ANIQSVLPSPSACASAGTNCNLISWGNTLSSGTNYPDLQLCNPALTDQSNSYPGEIT 178    | Db 119 ANIQSVLPSPSACASAGTNCNLISWGNTLSSGTNYPDLQLCNPALTDQSNSYPGEIT 178                                                                                                                                                          |                                                            |
| Qy                                                             | 181                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | EMNICACGPLEGGDSDGGPVGVCNGELOGTVSWGKGCAQKNPGVITKVCNFVDWIQ 240      | Qy 181 EMNICACGPLEGGDSDGGPVGVCNGELOGTVSWGKGCAQKNPGVITKVCNFVDWIQ 240     | Qy 181 EMNICACGPLEGGDSDGGPVGVCNGELOGTVSWGKGCAQKNPGVITKVCNFVDWIQ 240                                                                                                                                                           |                                                            |
| Db                                                             | 179                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | KMFECAGFLAGGDSDQGDSSGGPVYCNGLQGVYWSWGHGCAQKNPGVITKVCNFVDWIQ 238   | Db 179 KMFECAGFLAGGDSDQGDSSGGPVYCNGLQGVYWSWGHGCAQKNPGVITKVCNFVDWIQ 238  | Db 179 KMFECAGFLAGGDSDQGDSSGGPVYCNGLQGVYWSWGHGCAQKNPGVITKVCNFVDWIQ 238                                                                                                                                                        |                                                            |
| AC                                                             | X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | TRY1_XENLA STANDARD; PRT; 244 AA.                                 | AC ID TRY1_CANFA STANDARD; PRT; 246 AA.                                 | AC ID TRY1_CANFA STANDARD; PRT; 246 AA.                                                                                                                                                                                       |                                                            |
| PT                                                             | TRY2_XENLA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| AC                                                             | P7052;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| DT                                                             | 01-NOV-1997 (Rel. 35; Created)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                   |                                                                         | Qy 241 STAAN 246                                                                                                                                                                                                              |                                                            |
| DT                                                             | 01-NOV-1997 (Rel. 35; Last sequence update)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                   |                                                                         | Db 239 STISSN 244                                                                                                                                                                                                             |                                                            |
| DE                                                             | 15-JUN-2002 (Rel. 41; Last annotation update)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| DE                                                             | TRYPSIN precursor (EC 3.4.21.4)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| OS                                                             | Xenopus laevis (African clawed frog)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| OC                                                             | Amphibia; Batracia; Chordata; Craniata; Vertebrata; Euteleostomi;                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| OC                                                             | Xenopoda; Xenopus; Anura; Mesobatrachia; Pipidae; Xenopidae.                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| CC                                                             | XENOPODINAES; Xenopus; Xenopus.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                                                                         |                                                                                                                                                                                                                               |                                                            |
| CC                                                             | [1] RN SEQUENCE FROM N.A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                                                                   |                                                                         | OX NCBI_TAXID=9615;                                                                                                                                                                                                           |                                                            |
| RP                                                             | Wang K., Lyle L., Gan L., Hood L.E.; Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                   |                                                                         | RN SEQUENCE FROM N.A.                                                                                                                                                                                                         |                                                            |
| RA                                                             | 1-CATALYTIC ACTIVITY: Preferential cleavage: Arg -  -Xaa, Lys -  -Xaa.                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                   |                                                                         | RP SEQUENCE FROM N.A.                                                                                                                                                                                                         |                                                            |
| CC                                                             | -I- SUBCELLULAR LOCATION: Extracellular.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                   |                                                                         | RP MEDLINE=86284628; PubMed=3841794;                                                                                                                                                                                          |                                                            |
| CC                                                             | -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                   |                                                                         | RA Laforge K.S., Scheele G.;                                                                                                                                                                                                  |                                                            |
| CC                                                             | CC MEROPS: S01-258; -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                   |                                                                         | RA Pinsky S.D., Laforge K.S., Scheele G.;                                                                                                                                                                                     |                                                            |
| DR                                                             | InterPro; IPR00314; Chymotrypsin.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                   |                                                                         | DE "Differential regulation of trypsinogen mRNA translation: full-length mRNA sequences encoding two oppositely charged trypsinogen isoenzymes in the dog pancreas."                                                          |                                                            |
| DR                                                             | InterPro; IPR00254; Ser_Protease_Try.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                   |                                                                         | RT Mol. Cell. Biol. 5:2669-2676(1985).                                                                                                                                                                                        |                                                            |
| DR                                                             | Pfam: PF00089; trypsin_1_psipedia.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                   |                                                                         | CC -I- SUBCELLULAR LOCATION: Extracellular.                                                                                                                                                                                   |                                                            |
| DR                                                             | PRINTS; PRO0722; CHYNOTRYPSIN.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                   |                                                                         | CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI.                                                                                                                                                                            |                                                            |
| CC                                                             | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). |                                                                   |                                                                         | CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its |                                                            |

CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE AND FOR COMMERCIAL  
 CC ENTITIES REQUIRES A LICENSE AGREEMENT. (SEE [HTTP://WWW.ISB-SIB.CH/ANNOUNCE/](http://www.isb-sib.ch/announce/)  
 CC OR SEND AN EMAIL TO LICENSE@ISB-SIB.CH).

---

CC EMBL; M11590; AAA30900.1;  
 DR PIR; B26273; TRGCG.  
 DR MEROP; S01.151;  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser\_Protease\_TRY.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM0020; TRYSPC; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00155; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine proteases; Digestion; Pancreas; zymogen; Signal;  
 KW Multigene family.

SEQUENCE 1 15 SIGNAL 1 15 ACTIVATION PEPTIDE.  
 T PROSP 16 23 TRYPSIN, CATIONIC.  
 T CHAIN 24 246 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 T ACT\_SITE 63 63 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 T ACT\_SITE 107 107 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 T ACT\_SITE 200 200 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 T DISULFID 30 160 BY SIMILARITY.  
 T DISULFID 48 64 BY SIMILARITY.  
 T DISULFID 132 233 BY SIMILARITY.  
 T DISULFID 139 206 BY SIMILARITY.  
 T DISULFID 171 185 BY SIMILARITY.  
 T DISULFID 196 220 BY SIMILARITY.  
 T SITE 194 194 REQUIRED FOR SPECIFICITY (BY SIMILARITY).  
 SQ SEQUENCE 246 AA; 26170 MW; E9EAA1DE2391BBBB C9C64;

Query Match 77.3%; Score 1026; DB 1; Length 246;  
 Best Local Similarity 75.2%; Pred No. 3.5e-83; Indels 0; Gaps 0;  
 Matches 185; Conservative 23; Mismatches 38; Delins 0;

OY 1 MNPLLTIAFAAVATPTDDDKIVGGTCEENSVPQVQLSLNAGYHFGCGSLISDQWVS 60  
 Db 1 MKTFIFLALLGATVAFPIDDKIVGGTCSRNSVPQVQLSGYHFCCGSILNSDWVS 60  
 QY 61 AAHCYKSRIQVRGLGENIDLEQNEQITSAKVIIRHYNWILDNDIMLKLSPPAVLN 120  
 Db 61 AAHCYKSRIQVRGLGENITASSEGGEFINAKITIIRPYRNATIDIMLKLSPPAVLN 120  
 QY 121 ARVATSLPRCAAPGTCQCLISGRGNLTLSSGTNPPELQDAPITLQAQEASYPGQIT 180  
 , 121 SRVSATLPKSCPAAGTQCLISGWGNTQSICQNYPQQIS 180  
 QY 181 ENMIGAGFLEGGKDSQGDSCGPVCNGELOGIVSNGYGCRAQNKPGVYTRCNFWDWIQ 240  
 Db 181 SNMMCIGYMEGKDSQGDSSGPVYVNCNGELQGVSVPRVKCYVSQWI 240  
 QY 241 STIAAN 246  
 Db 241 QTIAN 246

RESULT 14

ID TRY1\_BOVIN STANDARD; PRT; 243 AA.

AC P00760; 21-JUL-1986 (Rel. 01, Created)

DT 01-NOV-1987 (Rel. 35, Last sequence update)

DE Trypsinogen, cationic precursor (EC 3.4.21.4) (Beta-trypsin)

DE (Fragment)

OS Bos taurus (Bovine)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;

OC Bovidae; Bovinae; Bos.

NCBI\_TaxID=9913;

[1]  
 RN SEQUENCE FROM N.A.  
 RP  
 RC  
 RA Okajima T.; Maniwa M.; Nagao S.; Fujikawa H.; Goto S.;  
 RL Submitted (Oct-1994) to the EMBL/GenBank/DBJ databases.

[2]  
 RN SEQUENCE OF 15-243, AND DISULFIDE BONDS.

RP MEDLINE-67163848; PubMed-5967094;  
 RX  
 RA Mikes O.; Holleyovský V.; Tomášek V.; Šorm F.;  
 RT "Covalent structure of bovine trypsinogen. The position of the  
 remaining amides";  
 RL Biochem. Biophys. Res. Commun. 24:346-352(1966).

[3]  
 RN REVISIONS.

RP MEDLINE-72035053; PubMed-4399051;  
 RX  
 RA Hartley B.S.;  
 RT "Homologues in serine proteinases.";  
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:77-87(1970).

[4]  
 RN REVIEWS.

RP MEDLINE-75146445; PubMed-1092332;  
 RX  
 RA Titani K.; Ericsson L.H.; Neurath H.; Walsh K.A.;  
 RT "Amino acid sequence of dogfish trypsin.";  
 RL Biochemistry 14:1358-1366(1975).

[5]  
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF CALCIUM-BINDING SITE.

RP MEDLINE-76072097; PubMed-512;  
 RX  
 RA Boden W.; Schwager P.;  
 RT "The refined crystal structure of bovine beta-trypsin at 1.8-A resolution. II. Crystallographic refinement, calcium binding site, benzamidine binding site and active site at pH 7.0.";  
 RT J. Mol. Biol. 98:693-717(1975).

[6]  
 RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RP MEDLINE-77112431; PubMed-556951;  
 RX  
 RA Kossiakoff A.A.; Chambers J.L.; Kay L.M.; Stroud R.M.;  
 RT "Structure of bovine trypsinogen at 1.9-A resolution.";  
 RL Biochemistry 16:654-664(1977).

[7]  
 RN DISULFIDE BONDS.

RP MEDLINE-66079271; PubMed-5892911;  
 RX  
 RA Keufflein D.L.;  
 RT "The disulfide bridges of trypsin.";  
 RL J. Mol. Biol. 12:929-932(1965).

[8]  
 CC -CATALYTIC ACTIVITY: Extracellular.

CC -SUBCELLULAR LOCATION: Extracellular.  
 CC -TISSUE/SPECIFICITY: SYNTHESIZED IN THE ACINAR CELLS OF THE  
 CC PANCREAS.

CC -PTM: AUTOCATALYTIC CLEAVAGE AFTER LYS-20 LEADS TO BETA-TRYPSIN BY  
 CC RELEASING A TERMINAL HEXAPEPTIDE. SUBSEQUENT CLEAVAGE AFTER  
 CC LYS-145 LEADS TO ALPHA-TRYPSIN. FURTHER CLEAVAGE AFTER LYS-190  
 CC YIELDS PSEUDOTRYPSIN. A CLEAVAGE MAY ALSO OCCUR AFTER ARG-119.  
 CC -SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -DATABASE: NAME=Worthington enzyme manual;

CC WWW="http://www.worthington-biochem.com/manual/T/TRY.html".

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

CC DR EMBL; D38507; BAA07516.1; -

CC DR PIR; A00946; TRBOTR.

CC DR PDB; 2TGA; 09-APR-85.

CC DR PDB; 1TGB; 14-MAR-85.

CC DR PDB; 1TGD; 09-APR-85.

CC DR PDB; 2TGN; 22-OCT-84.

CC DR PDB; 2TGP; 14-MAR-85.

CC DR PDB; 1TGS; 14-MAR-85.

|          |                                                                   |                      |     |                                |
|----------|-------------------------------------------------------------------|----------------------|-----|--------------------------------|
|          |                                                                   |                      |     | 243                            |
| DR       | ITGT; 15-OCT-90.                                                  | CHAIN                | 146 |                                |
| DR       | PDB; 2PST; 09-APR-85.                                             | CA_BIND              | 72  | 82                             |
| DR       | PDB; 1TLD; 15-OCT-92.                                             | BINDING              | 191 | 192                            |
| DR       | PDB; 2TLD; 15-JUL-92.                                             | FT BINDING           | 194 | 195                            |
| DR       | PDB; 1TPA; 14-MAR-85.                                             | FT BINDING           | 194 | 195                            |
| DR       | PDB; 2TPA; 15-JUL-91.                                             | FT BINDING           | 197 | 197                            |
| DR       | PDB; 3TPA; 14-MAR-85.                                             | FT DISULFID          | 27  | 157                            |
| DR       | PDB; 4TPA; 08-NOV-85.                                             | FT DISULFID          | 45  | 61                             |
| DR       | PDB; 1TPO; 14-MAR-85.                                             | FT DISULFID          | 129 | 230                            |
| DR       | PDB; 1TPP; 16-APR-87.                                             | FT DISULFID          | 136 | 203                            |
| DR       | PDB; 1NTP; 15-APR-91.                                             | FT DISULFID          | 168 | 182                            |
| DR       | PDB; 3PTB; 14-MAR-85.                                             | CHARGE RELAY SYSTEM. |     |                                |
| DR       | PDB; 2PTC; 14-MAR-85.                                             | ACT SITE             | 60  | 60                             |
| DR       | PDB; 2PTN; 15-APR-91.                                             | ACT_SITE             | 104 | 104                            |
| DR       | PDB; 3PTN; 09-APR-85.                                             | ACT_SITE             | 197 | 197                            |
| DR       | PDB; 1ISMF; 31-JUL-94.                                            | SITE                 | 191 | 191                            |
| DR       | PDB; 1TNG; 30-NOV-94.                                             | STRAND               | 22  | 22                             |
| DR       | PDB; 1TNH; 30-NOV-94.                                             | STRAND               | 25  | 26                             |
| DR       | PDB; 1TNJ; 30-NOV-94.                                             | TURN                 | 29  | 30                             |
| DR       | PDB; 1TNP; 31-JAN-94.                                             | STRAND               | 33  | 34                             |
| DR       | PDB; 1IPPE; 31-JAN-94.                                            | STRAND               | 35  | 39                             |
| DR       | PDB; 1IPPH; 31-JAN-94.                                            | STRAND               | 43  | 51                             |
| DR       | PDB; 1TNU; 30-NOV-94.                                             | TURN                 | 52  | 53                             |
| DR       | PDB; 1BTP; 29-NOV-96.                                             | STRAND               | 54  | 57                             |
| DR       | PDB; 1BTW; 15-OCT-95.                                             | HELIX                | 59  | 61                             |
| DR       | PDB; 1BTX; 15-OCT-95.                                             | STRAND               | 67  | 70                             |
| DR       | PDB; 1BTY; 15-OCT-95.                                             | TURN                 | 75  | 76                             |
| DR       | PDB; 1BTZ; 15-OCT-95.                                             | STRAND               | 83  | 92                             |
| DR       | PDB; 1IJRS; 14-OCT-96.                                            | TURN                 | 94  | 95                             |
| DR       | PDB; 1JRT; 14-OCT-96.                                             | STRAND               | 97  | 97                             |
| DR       | PDB; 1MAX; 14-OCT-96.                                             | TURN                 | 98  | 101                            |
| DR       | PDB; 1MAY; 14-OCT-96.                                             | STRAND               | 102 | 102                            |
| DR       | PDB; 1MTS; 20-AUG-97.                                             | TURN                 | 103 | 103                            |
| DR       | PDB; 1MTU; 12-NOV-97.                                             | STRAND               | 106 | 110                            |
| DR       | PDB; 1MTV; 12-NOV-97.                                             | STRAND               | 124 | 124                            |
| DR       | PDB; 1MTW; 12-NOV-97.                                             | TURN                 | 132 | 133                            |
| DR       | PDB; 1TPS; 26-JAN-95.                                             | STRAND               | 135 | 140                            |
| DR       | PDB; 1TYN; 26-JAN-95.                                             | STRAND               | 156 | 162                            |
| DR       | PDB; 1TAW; 24-JUN-97.                                             | HELIX                | 165 | 171                            |
| DR       | PDB; 1AQ7; 25-FEB-98.                                             | TURN                 | 173 | 174                            |
| DR       | PDB; 1TIO; 23-SEP-98.                                             | TURN                 | 178 | 179                            |
| DR       | PDB; 2TIO; 30-SEP-98.                                             | STRAND               | 180 | 183                            |
| DR       | PDB; 1AUJ; 16-DEC-98.                                             | TURN                 | 186 | 187                            |
| DR       | PDB; 1XUG; 16-DEC-98.                                             |                      |     |                                |
| DR       | PDB; 1XUH; 11-NOV-98.                                             |                      |     |                                |
| DR       | PDB; 1XUI; 11-NOV-98.                                             |                      |     |                                |
| DR       | PDB; 1XUJ; 11-NOV-98.                                             |                      |     |                                |
| DR       | PDB; 1XUK; 11-NOV-98.                                             |                      |     |                                |
| DR       | PDB; 1AUJ; 14-OCT-98.                                             |                      |     |                                |
| DR       | PDB; 1AZ8; 13-JAN-99.                                             |                      |     |                                |
| DR       | PDB; 1BJU; 13-JAN-99.                                             |                      |     |                                |
| DR       | PDB; 1BJV; 13-JAN-99.                                             |                      |     |                                |
| DR       | PDB; 2BZA; 23-MAR-99.                                             |                      |     |                                |
| DR       | PDB; 1CE5; 23-MAR-99.                                             |                      |     |                                |
| DR       | MEOPS; S01.15;                                                    |                      |     |                                |
| DR       | Intertro; IPR001254; Ser_protease_Try.                            |                      |     |                                |
| PFam;    | PF00089; trypsin; 1.                                              |                      |     |                                |
| SMART;   | SM00020; Tryp_Spc; 1.                                             |                      |     |                                |
| PROSITE; | PS50240; TRYPSIN_DOM; 1.                                          |                      |     |                                |
| PROSITE; | PS00134; TRYPSIN_HIS; 1.                                          |                      |     |                                |
| KW       | Hydrolase; Serine protease; Digestion; Pancreas; zymogen; Signal; |                      |     |                                |
| KW       | 3D-structure.                                                     |                      |     |                                |
| FT       | NON_TER                                                           | 1                    | 1   | RESULT 15                      |
| FT       | SIGNAL                                                            | <1                   | 14  | TRY1_XENLA                     |
| FT       | PROPEP                                                            | 15                   | 20  | ACTIVATION PEPTIDE.            |
| FT       | CHAIN                                                             | 21                   | 243 | TRYPSIN, CATIONIC.             |
| FT       | CHAIN                                                             | 21                   | 145 | ALPHA-TRYPSIN CHAIN 1.         |
| AC       |                                                                   |                      | ;   | 243 AA.                        |
| DT       |                                                                   |                      |     | 01-FEB-1991 (Rel. 17, Created) |

01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE TRYPSIN PRECURSOR (EC 3.4.21.4).  
 OS Xenopus laevis (African Clawed frog).  
 OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Xenopidae; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 TISSUE=Pancreas;  
 RX MEDLINE=9100725; PubMed=2210372;  
 RA Shi Y.B.; Brown D.D.;  
 RT "developmental and thyroid hormone-dependent regulation of pancreatic genes in Xenopus laevis".  
 RL Genes Dev. 4:1107-1113(1990).  
 CC -|- CATALYTIC ACTIVITY: preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.  
 CC -|- SUBCELLULAR LOCATION: Extracellular.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDE FAMILY S1.  
 C  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions is allowed as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).  
 C  
 EMBL; X53458; CAA31538.1; -.  
 DR PIR: A35871; A35871; -.  
 DR HSSP; P00763; IDPO.  
 DR MEROPS; S01.151; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001324; Serine\_protease\_TRY.  
 DR Pfam; PF0089; trypsin\_1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYPSIN\_SPC\_1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM\_1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER\_1.  
 KW Hydrolase; Serine Protease; Digestion; Zymogen; Signal;  
 KW Multigene family.  
 FT SIGNAL 1 15 BY SIMILARITY.  
 FT PROPEP 16 20 BY SIMILARITY.  
 FT CHAIN 21 243 TRYPSIN.  
 FT ACT\_SITE 60 60 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 104 104 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 T ACT SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 F DISULFID 27 157 BY SIMILARITY.  
 FT DISULFID 45 61 BY SIMILARITY.  
 FT DISULFID 129 230 BY SIMILARITY.  
 FT DISULFID 136 203 BY SIMILARITY.  
 FT DISULFID 168 182 BY SIMILARITY.  
 FT DISULFID 193 217 BY SIMILARITY.  
 FT SITE 191 191 REQUIRED FOR SPECIFICITY (BY SIMILARITY).  
 SQ 243 AA; 25492 MW; C5B8345AB3F8031 CRC64;

Query Match 74.8%; Score 992.5; DB 1; Length 243;  
 Best Local Similarity 73.6%; Pred. No. 3.1e-80;  
 Matches 181; Conservative 29; Mismatches 33; Indels 3; Gaps 1;

Qy 1 MNPLITLAFGLAAVATPDDDKIVGGTCCEENSYPPVQLNAGYHFCGGSLLISDOWWVS 60  
 DB 1 MKFLLCLVGIAAAAF--DDDKitGGCAKSSPPYIVSLNSGYHFCGGSLLTNQWWVS 57

Qy 61 AAHCYKSRQVRLGEYNIDYLENEQFINSAKVIRPHNNWILDNDIMIKLSSPAVLN 120  
 DB 58 AAHCYKASQVRLGEHNIAUSEQFISSKVIRSGNSYTLDDMIGIKLSSPAVLN 117

Qy 121 ARVATISLPACAAAGTQQLCISGNTLSSGTNTPELOCLADPLTAQCEASYPGQT 180  
 DB 118 AAVTVPVPLPSGCCSAGTSCSLISGNTLSNGSNPDLQCLNAPLNAQNSAYPGETI 177

GanCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:14:41 ; Search time 79 Seconds  
(without alignments)  
644.223 Million cell updates/sec

Title: US-09-762-277A-1  
Perfect score: 1327  
Sequence: 1 MNPLTLAFLGAAVATPTDD.....VYTKVCNFVDWIGSTIAANS 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues.

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21;\*

- 1: sp\_archea;\*
- 2: sp\_bacteria;\*
- 3: sp\_fungi;\*
- 4: sp\_human;\*
- 5: sp\_invertebrate;\*
- 6: sp\_mammal;\*
- 8: sp\_organelle;\*
- 9: sp\_phage;\*
- 10: sp\_plant;\*
- 11: sp\_rabbit;\*
- 12: sp\_virus;\*
- 13: sp\_vertebrate;\*
- 14: sp\_unclassified;\*
- 15: sp\_virus;\*
- 16: sp\_bacteriap;\*
- 17: sp\_archeap;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query | Match | Length | DB ID  | Description         |
|------------|--------|-------|-------|--------|--------|---------------------|
| 1          | 1119   | 84.3  | 246   | 11     | Q9QK9  | Q9qk9 mus musculu   |
| 2          | 1118   | 84.3  | 246   | 11     | Q9R07  | Q9r07 mus musculu   |
| 3          | 1099   | 82.8  | 246   | 11     | Q9Z1R9 | Q9z1r9 mus musculu  |
| 4          | 1063.5 | 60.1  | 247   | 11     | Q9CPN9 | Q9cpn9 mus musculu  |
| 5          | 1056.5 | 79.6  | 247   | 11     | Q9D7Y7 | Q9d7y7 mus musculu  |
| 6          | 970.5  | 73.1  | 247   | 11     | Q98T07 | Q98t07 mus musculu  |
| 7          | 906    | 68.3  | 241   | 13     | Q98TG9 | Q98tg9 engraulis j  |
| 8          | 899.5  | 67.8  | 237   | 13     | Q915L5 | Q915l5 fugu rubrip  |
| 9          | 897    | 67.6  | 242   | 13     | Q93266 | Q93266 pseudopleur  |
| 10         | 887    | 66.8  | 240   | 13     | Q98TH0 | Q98th0 engraulis j  |
| 11         | 879    | 66.2  | 242   | 13     | Q9WT07 | Q9w7q7 paralichthys |
| 12         | 872    | 65.7  | 244   | 13     | Q8QG93 | Q8qg9 anguilla ja   |
| 13         | 871.5  | 65.7  | 238   | 13     | Q9WT06 | Q9w7q6 paralichthys |
| 14         | 838.5  | 64.4  | 242   | 13     | Q92099 | Q92099 paranotothe  |
| 15         | 838.5  | 63.2  | 247   | 13     | Q42608 | Q42608 petromyzon   |
| 16         | 837.5  | 63.1  | 247   | 13     | Q9pt51 | Q9pt51 agkistrodon  |

## ALIGNMENTS

RESULT 1  
Q9QK9  
ID Q9QK9,  
AC Q9QK9,  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE TSP4 (0910001B19RK Protein (trypsinogen 9).  
GN TC OR 0910001B19RK OR TRYPSINOGEN.  
OS Mus musculus (Mouse).  
OU Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Etheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.  
RN NCBI\_TAXID=10090;  
RN [1]  
RN SEQUENCE FROM N.A.  
RN TISSUE=TESTIS;  
RX MEDLINE=99436155; PubMed=10506205;  
RA Ohmura K., Kohno N., Kobayashi Y., Yamagata K., Sato S.,  
RA Kashiwabara S., Baba T.;  
RT "A homologue of pancreatic trypsin is localized in the acrosome of  
RT mammalian sperm and is released during acrosome reaction.";  
RL J. Biol. Chem. 274:2926-29432(1999).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
RC MEDLINE=21056660; PubMed=11217811;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Miyosawa H., Kondo H., Saito R.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T.,  
RA Kadota K., Matsuda H.N., Ashburner M., Battalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Scabili F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

|                                  |                                                                                   |
|----------------------------------|-----------------------------------------------------------------------------------|
| RA                               | Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,                      |
| RA                               | Sasaki H., Sato K., Schoenbach C., Saya T., Shibusawa Y., Storch K.-F.,           |
| RA                               | Suzuki H., Toyooka K., Wang K.-H., Weitz C., Whittaker C., Wilming L.,            |
| RA                               | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,                |
| RA                               | "Functional annotation of a full-length mouse cDNA collection.";                  |
| RT                               | Nature 409:685-690(2001).                                                         |
| RN                               | [3]                                                                               |
| RP                               | SEQUENCE FROM N.A.                                                                |
| RC                               | STRAIN=129;                                                                       |
| RC                               | SEQUENCE FROM N.A.                                                                |
| RC                               | STRAIN=C57BL/6J; TISSUE=STOMACH, AND SPLEEN;                                      |
| RC                               | MEDLINE=21085660; PubMed=11217851;                                                |
| RX                               | "Differential transcriptional regulation of individual TCR Vbeta                  |
| RT                               | segments before gene rearrangement.;"                                             |
| RL                               | J. Immunol. 166:1771-1780(2001).                                                  |
| CC                               | -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE                 |
| DR                               | TRYPsin FAMILY.                                                                   |
| EMBL; AB009661; BAA85187..1; ..  |                                                                                   |
| EMBL; AB010731; BAA74760..1; ..  |                                                                                   |
| EMBL; AK007843; BAA25500..1; ..  |                                                                                   |
| EMBL; AE000664; AAB69057..1; ..  |                                                                                   |
| HSPP; P00763; 1POPO.             |                                                                                   |
| DR                               | MEROPS; S01_058; ..                                                               |
| MDG; MGI:1913350; 0910001B19Rik. |                                                                                   |
| DR                               | InterPro; IPR001314; Chymotrypsin.                                                |
| DR                               | InterPro; IPR001254; Ser_protease_Try.                                            |
| DR                               | Pfam; PF00089; trypsin; 1.                                                        |
| DR                               | PRINTS; PR0722; CHYMO TRYPSIN.                                                    |
| DR                               | SMART; SM00020; TRYP_SPC; 1.                                                      |
| DR                               | PROSITE; PS50240; TRIPSIN_DOM; 1.                                                 |
| DR                               | PROSITE; PS00134; TRIPSIN_HIS; UNKNOWN_1.                                         |
| DR                               | PROSITE; PS00135; TRIPSIN_SER; 1.                                                 |
| RW                               | Hydrolease; Serine Protease.                                                      |
| SQ                               | SEQUENCE 246 AA; 26277 MW; 915C92353EG42809 CRC64;                                |
|                                  | [3]                                                                               |
| Query Match                      | 84.3%; Score 1119; DB 11; Length 246;                                             |
| Best Local Similarity            | 81.7%; Pred. No. 1.2e-95;                                                         |
| Matches                          | 201; Conservative 22; Mismatches 23; Indels 0; Gaps 0;                            |
| OY                               | 1 MNPLLIAFLGLAVATPDDDKIVGGTYCEENSYPTOVSLSNAGYHFGGSLISDNQVVS 60                    |
| Db                               | 1 MNSLLFLALGVAVAFPVDDDKIVGGTYCRENISIPTOVSLSNAGYHFGGSLISDNQVVS 60                  |
| OY                               | 61 AAHCYKSRQVRLGEYNIDVLESGNEQFTINSAKVIRHPNYSWILDMLIKLISSPAVLN 120                 |
| Db                               | 61 AAHCYKTRIQVNLGEHNINLVEGNEQFNSAKTIKHFNENSPTLNNDMLIKLISSPAVLN 120                |
| OY                               | 61 AAHCYKSRQVRLGEYNIDVLESGNEQFTINSAKVIRHPNYSWILDMLIKLISSPAVLN 120                 |
| Db                               | 61 AAHCYKTRIQVNLGEHNINLVEGNEQFNSAKTIKHFNENSPTLNNDMLIKLISSPAVLN 120                |
| OY                               | 121 ARVATISPLRACAPGTQCLISGWNTLSSGTNPYELQCLADPLTQAQCEASYPGQIT 180                  |
| Db                               | 121 ARVATVPLPSSCAPAGTQCLISGWNTLSSGTNPYELQCLADPLTQAQCEASYPGQIT 180                 |
| OY                               | 181 ENMICAFCGLEKGDSCQDGGFPVYCNGELOQTSVSYGYCAQKNPKGYTQVCFNFVWIK 240                |
| Db                               | 181 NNMCYQGLEGRDSCQDGGFPVVCNCQLGQTVSVSYGCAKLDNPQGYTTKVCNVWIK 240                  |
| OY                               | 241 STIAAN 246                                                                    |
| Db                               | 241 DTAAN 246                                                                     |
| RESULT 2                         |                                                                                   |
| O9R0T7                           | PRELIMINARY; PRT; 246 AA.                                                         |
| ID                               | O9R0T7                                                                            |
| AC                               | O9R0T7;                                                                           |
| DT                               | 01-MAY-2000 (TRIMBLrel. 13, Created)                                              |
| DT                               | 01-MAY-2000 (TRIMBLrel. 13, Last sequence update)                                 |
| DT                               | 01-JUN-2002 (TRIMBLrel. 21, Last annotation update)                               |
| DE                               | Pancretic trypsin (0910001B19Rik protein) (Trypsinogen 8).                        |
| GN                               | TD OR 0910001B19Rik OR TRYPsinOgen.                                               |
| OS                               | Mus musculus (Mouse)                                                              |
| OC                               | Mammalia; Euteleostomi; Vertebrata; Craniata; Sciurognathi; Muridae; Murinae; Mus |
| OX                               | NCBI_TaxID=10090;                                                                 |
| Query Match                      | 84.3%; Score 1118; DB 11; Length 246;                                             |
| Best Local Similarity            | 82.1%; Mismatches 21; Indels 0; Gaps 0;                                           |
| Matches                          | 202; Conservative 21; Mismatches 23; Indels 0; Gaps 0;                            |
| OY                               | 1 MNPLLIAFLGLAVATPDDDKIVGGTYCEENSYPTOVSLSNAGYHFGGSLISDNQVVS 60                    |
| Db                               | 1 MRLLFLALGVAVAFPVDDDKIVGGTYCRENISIPTOVSLSNAGYHFGGSLISDNQVVS 60                   |
| OY                               | 61 AAHCYKSRQVRLGEYNIDVLESGNEQFTINSAKVIRHPNYSWILDMLIKLISSPAVLN 120                 |
| Db                               | 61 AAHCYKTRIQVNLGEHNINLVEGNEQFNSAKTIKHFNENSPTLNNDMLIKLISSPAVLN 120                |



| SO                          | SEQUENCE                                                                          | 247 AA: | 264/22 MW;         | B8C5767B182D9AAD CRC64; |
|-----------------------------|-----------------------------------------------------------------------------------|---------|--------------------|-------------------------|
| Query Match                 | Score 1063.5;                                                                     | DB 11;  | Length 247;        |                         |
| Best Local Similarity       | 80.1%;                                                                            |         | Pred. No. 1.7e-90; |                         |
| Matches 188; Conservative   | 76.1%;                                                                            |         | Mismatches 31;     | UNKNOWN_1.              |
| Matches 188; Mismatches 27; | Indels 1;                                                                         | Gaps 1; |                    |                         |
| Qy                          | 1 MNPLLILAFLAGAAVATP-TDDDKIVGGTCTEENSPVQVSINAGYHFCGGSLISDQNVY 59                  |         |                    |                         |
| Db                          | 1 SAACYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                           |         |                    |                         |
| Qy                          | 60 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Db                          | 61 SAACYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                          |         |                    |                         |
| Qy                          | 120 NARVATISLPRACAAAPGTCQCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179             |         |                    |                         |
| Db                          | 121 NSRVSFTALPRSCPSAGTRCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179               |         |                    |                         |
| Qy                          | 180 TENMICAGFLEGGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Db                          | 181 TSNMFCLGFLLEGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Qy                          | 240 QSTIAAN 246                                                                   |         |                    |                         |
| Db                          | 241 QOTIAAN 247                                                                   |         |                    |                         |
| RESULT 5                    |                                                                                   |         |                    |                         |
| Q9D7Y7                      | PRELIMINARY:                                                                      | PRT;    | 247 AA.            |                         |
| AC                          | Q9D7Y7;                                                                           |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Created)                                              |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Last sequence update)                                 |         |                    |                         |
| DT                          | 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                               |         |                    |                         |
| DE                          | 2210010C0Rik Protein.                                                             |         |                    |                         |
| GN                          | 2210010C0Rik.                                                                     |         |                    |                         |
| OS                          | Mus musculus (Mouse).                                                             |         |                    |                         |
| OC                          | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                 |         |                    |                         |
| OC                          | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.                |         |                    |                         |
| OX                          | NCBI_TaxID=10090;                                                                 |         |                    |                         |
| RN                          | [1]                                                                               |         |                    |                         |
| RP                          | SEQUENCE FROM N.A.                                                                |         |                    |                         |
| RC                          | STRAIN=C57BL/6J; TISSUE=STOMACH;                                                  |         |                    |                         |
| RX                          | MEDLINE=2085660; PubMed=11218751;                                                 |         |                    |                         |
| RA                          | Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,                 |         |                    |                         |
| RA                          | Akara A., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,                  |         |                    |                         |
| RA                          | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane K.,                  |         |                    |                         |
| RA                          | Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,                |         |                    |                         |
| RA                          | Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,                   |         |                    |                         |
| RA                          | Kuehl P., Lewis S., Nikaido I., Pesole G., Quackenbush J.,                        |         |                    |                         |
| RA                          | Schrimal L.M., Matsumoto Y., Suzuki R., Tomita M., Wagner L., Washio T.,          |         |                    |                         |
| RA                          | Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,                  |         |                    |                         |
| RA                          | Gustincich S., Hill D., Hofmann M., Hume D.A., Kanoya M., Lee N.H.,               |         |                    |                         |
| RA                          | Lyon P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,                  |         |                    |                         |
| RA                          | Nordoni H., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,                      |         |                    |                         |
| RA                          | Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,             |         |                    |                         |
| RA                          | Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,             |         |                    |                         |
| RA                          | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,                |         |                    |                         |
| RA                          | Hayashizaki Y.;                                                                   |         |                    |                         |
| RT                          | "Functional annotation of a full-length mouse cDNA collection.";                  |         |                    |                         |
| RL                          | Nature 409:685-690(2001).                                                         |         |                    |                         |
| CC                          | -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI; ALSO KNOWN AS THE TRYPsin FAMILY. |         |                    |                         |
| CC                          | HSSP: P000695; BAB25837.1; -                                                      |         |                    |                         |
| DR                          | MEIROS; S01.151; -                                                                |         |                    |                         |
| DR                          | HSSP: P000695; 1DP0.                                                              |         |                    |                         |
| DR                          | MGRD; MGII:1914623; 2210010C04R1K.                                                |         |                    |                         |
| DR                          | InterPro; IPR001314; Chymotrypsin.                                                |         |                    |                         |
| DR                          | InterPro; IPR001254; Ser_protease Try.                                            |         |                    |                         |
| DR                          | Pfam; PF00089; trypsin_1.                                                         |         |                    |                         |
| PRINTS;                     | PR00722; CHYMOTRYPSIN.                                                            |         |                    |                         |
| SMART;                      | SM00020; TRYSP_CPC_1.                                                             |         |                    |                         |
| PROSITE;                    | PS50240; TRYPSIN_DOM; 1.                                                          |         |                    |                         |
| PROSITE;                    | PS00134; TRYPSIN_HIS; UNKNOWN_1.                                                  |         |                    |                         |
| PROSITE;                    | PS00135; TRYPSIN_SER; 1.                                                          |         |                    |                         |
| KW                          | Hydrolase; Serine protease.                                                       |         |                    |                         |
| SEQUENCE                    | 247 AA; 264/07 MW;                                                                |         |                    |                         |
| Qy                          | Query Match Score 1056.5; DB 11; Length 247;                                      |         |                    |                         |
| Db                          | Best Local Similarity 75.7%; Pred. No. 7.7e-90;                                   |         |                    |                         |
| Qy                          | Best Local Similarity 75.7%; Mismatches 31; MisMatch 28;                          |         |                    |                         |
| Db                          | Indels 1; Gaps 1; Matches 187; Conservative 31; Mismatches 28;                    |         |                    |                         |
| Qy                          | Query Match Score 1056.5; DB 11; Length 247;                                      |         |                    |                         |
| Db                          | Best Local Similarity 75.7%; Pred. No. 7.7e-90;                                   |         |                    |                         |
| Qy                          | Best Local Similarity 75.7%; Mismatches 31; MisMatch 28;                          |         |                    |                         |
| Db                          | Indels 1; Gaps 1; Matches 187; Conservative 31; Mismatches 28;                    |         |                    |                         |
| Qy                          | 1 MNPLLILAFLAGAAVATP-TDDDKIVGGTCTEENSPVQVSINAGYHFCGGSLISDQNVY 59                  |         |                    |                         |
| Db                          | 1 MKTLILAFLAGAAVALPDDDKIVGGTCTORNALPQVSINAGYHFCGGSLINSQNVY 60                     |         |                    |                         |
| Qy                          | 60 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Db                          | 61 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Qy                          | 120 NARVATISLPRACAAAPGTCQCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179             |         |                    |                         |
| Db                          | 121 NSRVSFTALPRSCPSAGTRCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179               |         |                    |                         |
| Qy                          | 180 TENMICAGFLEGGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Db                          | 181 TSNMFCLGFLLEGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Qy                          | 240 QSTIAAN 246                                                                   |         |                    |                         |
| Db                          | 241 QOTIAAN 247                                                                   |         |                    |                         |
| Qy                          | 180 TENMICAGLEGGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                  |         |                    |                         |
| Db                          | 181 TSNMFCLGPLEGGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Qy                          | 240 QSTIAAN 246                                                                   |         |                    |                         |
| Db                          | 241 QOTIAAN 247                                                                   |         |                    |                         |
| Q9D7Y7                      | PRELIMINARY:                                                                      | PRT;    | 247 AA.            |                         |
| AC                          | Q9D7Y7;                                                                           |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Created)                                              |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Last sequence update)                                 |         |                    |                         |
| DT                          | 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                               |         |                    |                         |
| DE                          | 181009306Rik protein (trypsinogen 4).                                             |         |                    |                         |
| GN                          | 181009306Rik OR TRYPSINOGEN.                                                      |         |                    |                         |
| OS                          | Mus musculus (Mouse).                                                             |         |                    |                         |
| OC                          | Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;                          |         |                    |                         |
| OC                          | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.                |         |                    |                         |
| OX                          | NCBI_TaxID=10090;                                                                 |         |                    |                         |
| RN                          | [1]                                                                               |         |                    |                         |
| RP                          | SEQUENCE FROM N.A.                                                                |         |                    |                         |
| RC                          | STRAIN=C57BL/6J; TISSUE=Pancreas;                                                 |         |                    |                         |
| RX                          | MEDLINE=21085660; PubMed=11217851;                                                |         |                    |                         |
| RA                          | Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,                |         |                    |                         |
| RA                          | Akara A., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,                  |         |                    |                         |
| RA                          | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane K.,                  |         |                    |                         |
| RA                          | Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,                |         |                    |                         |
| RA                          | Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,                   |         |                    |                         |
| RA                          | Kuehl P., Lewis S., Nikaido I., Pesole G., Quackenbush J.,                        |         |                    |                         |
| RA                          | Schrimal L.M., Matsumoto Y., Suzuki R., Tomita M., Wagner L., Washio T.,          |         |                    |                         |
| RA                          | Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,                  |         |                    |                         |
| RA                          | Gustincich S., Hill D., Hofmann M., Hume D.A., Kanoya M., Lee N.H.,               |         |                    |                         |
| RA                          | Lyon P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,                  |         |                    |                         |
| RA                          | Nordoni H., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,                      |         |                    |                         |
| RA                          | Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,             |         |                    |                         |
| RA                          | Suzuki H., Toyo-oka K., Wang K.H., Whittaker C., Wilming L.,                      |         |                    |                         |
| RA                          | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,                |         |                    |                         |
| RA                          | Hayashizaki Y.;                                                                   |         |                    |                         |
| RT                          | "Functional annotation of a full-length mouse cDNA collection.";                  |         |                    |                         |
| RL                          | Nature 409:685-690(2001).                                                         |         |                    |                         |
| CC                          | -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI; ALSO KNOWN AS THE TRYPsin FAMILY. |         |                    |                         |
| CC                          | HSSP: P000695; BAB25837.1; -                                                      |         |                    |                         |
| DR                          | MEIROS; S01.151; -                                                                |         |                    |                         |
| DR                          | HSSP: P000695; 1DP0.                                                              |         |                    |                         |
| DR                          | MGRD; MGII:1914623; 2210010C04R1K.                                                |         |                    |                         |
| DR                          | InterPro; IPR001314; Chymotrypsin.                                                |         |                    |                         |
| DR                          | InterPro; IPR001254; Ser_protease Try.                                            |         |                    |                         |
| DR                          | Pfam; PF00089; trypsin_1.                                                         |         |                    |                         |
| PRINTS;                     | PR00722; CHYMOTRYPSIN.                                                            |         |                    |                         |
| SMART;                      | SM00020; TRYSP_CPC_1.                                                             |         |                    |                         |
| PROSITE;                    | PS50240; TRYPSIN_DOM; 1.                                                          |         |                    |                         |
| PROSITE;                    | PS00134; TRYPSIN_HIS; UNKNOWN_1.                                                  |         |                    |                         |
| PROSITE;                    | PS00135; TRYPSIN_SER; 1.                                                          |         |                    |                         |
| KW                          | Hydrolase; Serine protease.                                                       |         |                    |                         |
| SEQUENCE                    | 247 AA; 264/07 MW;                                                                |         |                    |                         |
| Qy                          | Query Match Score 1056.5; DB 11; Length 247;                                      |         |                    |                         |
| Db                          | Best Local Similarity 75.7%; Pred. No. 7.7e-90;                                   |         |                    |                         |
| Qy                          | Best Local Similarity 75.7%; Mismatches 31; MisMatch 28;                          |         |                    |                         |
| Db                          | Indels 1; Gaps 1; Matches 187; Conservative 31; Mismatches 28;                    |         |                    |                         |
| Qy                          | Query Match Score 1056.5; DB 11; Length 247;                                      |         |                    |                         |
| Db                          | Best Local Similarity 75.7%; Pred. No. 7.7e-90;                                   |         |                    |                         |
| Qy                          | Best Local Similarity 75.7%; Mismatches 31; MisMatch 28;                          |         |                    |                         |
| Db                          | Indels 1; Gaps 1; Matches 187; Conservative 31; Mismatches 28;                    |         |                    |                         |
| Qy                          | 1 MNPLLILAFLAGAAVATP-TDDDKIVGGTCTEENSPVQVSINAGYHFCGGSLISDQNVY 59                  |         |                    |                         |
| Db                          | 1 MKTLILAFLAGAAVALPDDDKIVGGTCTORNALPQVSINAGYHFCGGSLINSQNVY 60                     |         |                    |                         |
| Qy                          | 60 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Db                          | 61 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Qy                          | 120 NARVATISLPRACAAAPGTCQCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179             |         |                    |                         |
| Db                          | 121 NSRVSFTALPRSCPSAGTRCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179               |         |                    |                         |
| Qy                          | 180 TENMICAGFLEGGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Db                          | 181 TSNMFCLGFLLEGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Qy                          | 240 QSTIAAN 246                                                                   |         |                    |                         |
| Db                          | 241 QOTIAAN 247                                                                   |         |                    |                         |
| Q9D7Y7                      | PRELIMINARY:                                                                      | PRT;    | 247 AA.            |                         |
| AC                          | Q9D7Y7;                                                                           |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Created)                                              |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Last sequence update)                                 |         |                    |                         |
| DT                          | 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                               |         |                    |                         |
| DE                          | 181009306Rik protein (trypsinogen 4).                                             |         |                    |                         |
| GN                          | 181009306Rik OR TRYPSIN.                                                          |         |                    |                         |
| OS                          | Mus musculus (Mouse).                                                             |         |                    |                         |
| OC                          | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                 |         |                    |                         |
| OC                          | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.                |         |                    |                         |
| OX                          | NCBI_TaxID=10090;                                                                 |         |                    |                         |
| RN                          | [1]                                                                               |         |                    |                         |
| RP                          | SEQUENCE FROM N.A.                                                                |         |                    |                         |
| RC                          | STRAIN=C57BL/6J; TISSUE=STOMACH;                                                  |         |                    |                         |
| RX                          | MEDLINE=2085660; PubMed=11218751;                                                 |         |                    |                         |
| RA                          | Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,                |         |                    |                         |
| RA                          | Akara A., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,                  |         |                    |                         |
| RA                          | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane K.,                  |         |                    |                         |
| RA                          | Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,                |         |                    |                         |
| RA                          | Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,                   |         |                    |                         |
| RA                          | Kuehl P., Lewis S., Nikaido I., Pesole G., Quackenbush J.,                        |         |                    |                         |
| RA                          | Schrimal L.M., Matsumoto Y., Suzuki R., Tomita M., Wagner L., Washio T.,          |         |                    |                         |
| RA                          | Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,                  |         |                    |                         |
| RA                          | Gustincich S., Hill D., Hofmann M., Hume D.A., Kanoya M., Lee N.H.,               |         |                    |                         |
| RA                          | Lyon P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,                  |         |                    |                         |
| RA                          | Nordoni H., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,                      |         |                    |                         |
| RA                          | Sasaki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,             |         |                    |                         |
| RA                          | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,                |         |                    |                         |
| RA                          | Hayashizaki Y.;                                                                   |         |                    |                         |
| RT                          | "Functional annotation of a full-length mouse cDNA collection.";                  |         |                    |                         |
| RL                          | Nature 409:685-690(2001).                                                         |         |                    |                         |
| CC                          | -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI; ALSO KNOWN AS THE TRYPsin FAMILY. |         |                    |                         |
| CC                          | HSSP: P000695; BAB25837.1; -                                                      |         |                    |                         |
| DR                          | MEIROS; S01.151; -                                                                |         |                    |                         |
| DR                          | HSSP: P000695; 1DP0.                                                              |         |                    |                         |
| DR                          | MGRD; MGII:1914623; 2210010C04R1K.                                                |         |                    |                         |
| DR                          | InterPro; IPR001314; Chymotrypsin.                                                |         |                    |                         |
| DR                          | InterPro; IPR001254; Ser_protease Try.                                            |         |                    |                         |
| DR                          | Pfam; PF00089; trypsin_1.                                                         |         |                    |                         |
| PRINTS;                     | PR00722; CHYMOTRYPSIN.                                                            |         |                    |                         |
| SMART;                      | SM00020; TRYSP_CPC_1.                                                             |         |                    |                         |
| PROSITE;                    | PS50240; TRYPSIN_DOM; 1.                                                          |         |                    |                         |
| PROSITE;                    | PS00134; TRYPSIN_HIS; UNKNOWN_1.                                                  |         |                    |                         |
| PROSITE;                    | PS00135; TRYPSIN_SER; 1.                                                          |         |                    |                         |
| KW                          | Hydrolase; Serine protease.                                                       |         |                    |                         |
| SEQUENCE                    | 247 AA; 264/07 MW;                                                                |         |                    |                         |
| Qy                          | Query Match Score 1056.5; DB 11; Length 247;                                      |         |                    |                         |
| Db                          | Best Local Similarity 75.7%; Pred. No. 7.7e-90;                                   |         |                    |                         |
| Qy                          | Best Local Similarity 75.7%; Mismatches 31; MisMatch 28;                          |         |                    |                         |
| Db                          | Indels 1; Gaps 1; Matches 187; Conservative 31; Mismatches 28;                    |         |                    |                         |
| Qy                          | Query Match Score 1056.5; DB 11; Length 247;                                      |         |                    |                         |
| Db                          | Best Local Similarity 75.7%; Pred. No. 7.7e-90;                                   |         |                    |                         |
| Qy                          | Best Local Similarity 75.7%; Mismatches 31; MisMatch 28;                          |         |                    |                         |
| Db                          | Indels 1; Gaps 1; Matches 187; Conservative 31; Mismatches 28;                    |         |                    |                         |
| Qy                          | 1 MNPLLILAFLAGAAVATP-TDDDKIVGGTCTEENSPVQVSINAGYHFCGGSLISDQNVY 59                  |         |                    |                         |
| Db                          | 1 MKTLILAFLAGAAVALPDDDKIVGGTCTORNALPQVSINAGYHFCGGSLINSQNVY 60                     |         |                    |                         |
| Qy                          | 60 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Db                          | 61 SAAHCYKSRQIQLRGEQFTDAAKIKIRHNVNANTYNNDIMLILKTTAATL 120                         |         |                    |                         |
| Qy                          | 120 NARVATISLPRACAAAPGTCQCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179             |         |                    |                         |
| Db                          | 121 NSRVSFTALPRSCPSAGTRCLISGWNTLSSGTNPYELLOQCLADAPTQAOCEASYPGQI 179               |         |                    |                         |
| Qy                          | 180 TENMICAGFLEGGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Db                          | 181 TSNMFCLGFLLEGKDSCGDGGPQVYCNGLGEGIYSGCGAOKNPKGSVYTWCNFYDWI 239                 |         |                    |                         |
| Qy                          | 240 QSTIAAN 246                                                                   |         |                    |                         |
| Db                          | 241 QOTIAAN 247                                                                   |         |                    |                         |
| Q9D7Y7                      | PRELIMINARY:                                                                      | PRT;    | 247 AA.            |                         |
| AC                          | Q9D7Y7;                                                                           |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Created)                                              |         |                    |                         |
| DT                          | 01-JUN-2001 (TREMBLrel. 17, Last sequence update)                                 |         |                    |                         |
| DT                          | 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                               |         |                    |                         |
| DE                          | 181009306Rik protein (trypsinogen 4).                                             |         |                    |                         |
| GN                          | 181009306Rik OR TRYPSIN.                                                          |         |                    |                         |
| OS                          | Mus musculus (Mouse).                                                             |         |                    |                         |
| OC                          | Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;                          |         |                    |                         |
| OC                          | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.                |         |                    |                         |
| OX                          | NCBI_TaxID=10090;                                                                 |         |                    |                         |
| RN                          | [1]                                                                               |         |                    |                         |
| RP                          | SEQUENCE FROM N.A.                                                                |         |                    |                         |
| RC                          | STRAIN=C57BL/6J; TISSUE=Pancreas;                                                 |         |                    |                         |
| RX                          | MEDLINE=21085660; PubMed=11217851;                                                |         |                    |                         |
| RA                          | Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,                |         |                    |                         |
| RA                          | Akara A., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,                  |         |                    |                         |
| RA                          | Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane K.,                  |         |                    |                         |
| RA                          | Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,                |         |                    |                         |
| RA                          | Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,                   |         |                    |                         |
| RA                          | Kuehl P., Lewis S., Nikaido I., Pesole G., Quackenbush J.,                        |         |                    |                         |
| RA                          | Schrimal L.M., Matsumoto Y., Suzuki R., Tomita M., Wagner L., Washio T.,          |         |                    |                         |
| RA                          | Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,                  |         |                    |                         |
| RA                          | Gustincich S., Hill D., Hofmann M., Hume D.A., Kanoya M., Lee N.H.,               |         |                    |                         |
| RA                          | Lyon P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,                  |         |                    |                         |
| RA                          | Nordoni H., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,                      |         |                    |                         |
| RA                          | Sasaki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,             |         |                    |                         |
| RA                          | Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,                |         |                    |                         |
| RA                          | Hayashizaki Y.;                                                                   |         |                    |                         |
| RT                          | "Functional annotation of a full-length mouse cDNA collection.";                  |         |                    |                         |
| RL                          | Nature 409:685-690(2001).                                                         |         |                    |                         |
| CC                          | -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SI; ALSO KNOWN AS THE TRYPsin FAMILY. |         |                    |                         |
| CC                          | HSSP: P000695; BAB25837.1; -                                                      |         |                    |                         |
| DR                          | MEIROS; S01.151; -                                                                |         |                    |                         |
| DR                          | HSSP: P000695; 1DP0.                                                              |         |                    | </td                    |

RA MEDLINE=21103195; PubMed=11160223;  
 RA Chen F., Rowen L., Hood L., Rothenberg E.V.;  
 RT "Differential transcriptional regulation of individual TCR Vbeta  
 segments before gene rearrangement.";  
 RL J. Immunol. 166: 1771-1780 (2001).  
 CC -I SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AK007406; BA25018\_1; -.  
 DR EMBL; AE000663; AA869055\_1; -.  
 DR HSSP; P00763; 1DPO.  
 DR MEROPS; S01.129; -.  
 DR MGD; MGI:1920876; 1810009J06Rik.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001325; Ser\_Protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PRO00722; CHYMO TRYPSIN.  
 DR SMART; SM00020; tryp\_SPC; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR Hydrolase; Serine protease.  
 SQ SEQUENCE 241 AA; 26282 MW; FE362D39CAEBB2F6 CRC64;

Query Match 68.3%; Score 906; DB 13; Length 241;  
 Best Local Similarity 67.2%; Pred. No. 6.e-76;  
 Matches 164; Conservative 34; Mismatches 40; Indels 6; Gaps 2;

DR MNPLILIAFLGAAVATPDDDKIVGGTCEENSPVQVSLSNAGYHFCGSSLISDQWVVS 60  
 DR SMART; SM00020; TRYSPC; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR Hydrolase; Serine protease.  
 SQ SEQUENCE 247 AA; 26503 MW; EDF5FP7636833CTBC CRC64;

Query Match 73.1%; Score 970.5; DB 11; Length 247;  
 Best Local Similarity 70.4%; Pred. No. 7.e-82;  
 Matches 174; Conservative 31; Mismatches 41; Indels 1; Gaps 1;

DR MNPLILIAFLGAAVATPDDDKIVGGTCEENSPVQVSLSNAGY-HFCGSSLISDQWVV 59  
 DR MKIITFFFLGAVALPANSDDKIVGGTCEENSPVQVSLSNAGYHFCGSSLISDQWVL 60

DR SAACGYKSRQLQVRGEYNDVLEGNEQFINSAKVIRHPNTNSWLDIMLIKSPAVL 119  
 DR 61 SAACGYKSRQLQVRGEYNDVLEGNEQFIDSERKIRHPDINKTDVNDIMLIKSPAIL 120

DR 120 NARYATISLPRACAPGTOCLISGMGNTLSSGTNTNPELQLADPLTQAOCEASYPGQI 179  
 DR 121 NSQVSTVSLPSRCASTNAQCLVSGMGNTVSIGKTPALLQCLEAPVLSASSCKKSYPGQI 180

DR 180 TENMTCAGFLLEGKDSCQDSGGPPVYCNGELQGTYSWWGYCAQRNKPQVTTKVGNFDWI 239  
 DR 181 TSNNMFCLGFLEGGKDSCDGGPPVYCNGBQGTYSWWGSTCAMRKSPQVTTKVCNLYLSWI 240

DR 240 OSTIAAN 246  
 DR 241 QETMANN 247

RESULTS 7  
 ID Q98TG9 PRELIMINARY; PRT; 241 AA.  
 AC Q98TG9; PRELIMINARY;  
 DT 09-BR 2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE TRYPSinogen II.  
 GN ATRYII  
 OS Engraulis japonicus (Japanese anchovy).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;  
 OC Engraulidae.  
 OX NCBI\_TaxID=42892;

RN 1  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLYLIC CAECA;  
 RA Watabe S., Ahsan M.N.; Funabara D.;  
 RT "Anchovy trypsinogen manA.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.  
 CC -I SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AB041930; BAB40330\_1; -.  
 DR HSSP; P00763; 1DPO.  
 DR MEROPS; S01.258; -.



|                                                                                        |              |      |         |  |  |        |                                                                                                                   |
|----------------------------------------------------------------------------------------|--------------|------|---------|--|--|--------|-------------------------------------------------------------------------------------------------------------------|
| Q9W7Q7                                                                                 | PRELIMINARY; | PRT; | 242 AA. |  |  | OX     | NCBI_TaxID=7937;                                                                                                  |
| AC Q9W7Q7;                                                                             |              |      |         |  |  | RN [1] | SEQUENCE FROM N.A.                                                                                                |
| DT 01-NOV-1999 (TREMBLrel. 12, Created)                                                |              |      |         |  |  | RP     | RC TISSUE=PANCREAS;                                                                                               |
| DT 01-JUN-1999 (TREMBLrel. 12, Last sequence update)                                   |              |      |         |  |  | RA     | Kurokawa T., Suzuki T., Ohta H., Kogawa H., Tanaka H., Unuma T.;                                                  |
| DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                                 |              |      |         |  |  | RT     | "Expression of pancreatic enzyme genes during the early larval stage of Japanese eel, <i>Anguilla japonica</i> ." |
| DE Trypsinogen 1.                                                                      |              |      |         |  |  | RT     | Submitted (AUG-2001) to EMBL/GenBank/DDBJ databases.                                                              |
| OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                   |              |      |         |  |  | DR     | EMBL: AB07020; BAB8634.1; -.                                                                                      |
| OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;                  |              |      |         |  |  | KW     | Hydrolase.                                                                                                        |
| OC Acanthomorpha; Acanthopterygii; Percormorpha; Pleuronectiformes;                    |              |      |         |  |  | SQ     | SEQUENCE 244 AA; 26317 MW; 0EB3B68E8706D52D CRC64;                                                                |
| CC Pleuronectoidei; Paralichthyidae; Paralichthys.                                     |              |      |         |  |  |        |                                                                                                                   |
| OX NCBI_TaxID=8255;                                                                    |              |      |         |  |  |        |                                                                                                                   |
| RN                                                                                     |              |      |         |  |  |        |                                                                                                                   |
| RP                                                                                     |              |      |         |  |  |        |                                                                                                                   |
| RC                                                                                     |              |      |         |  |  |        |                                                                                                                   |
| SEQUENCE FROM N.A.                                                                     |              |      |         |  |  |        |                                                                                                                   |
| RA                                                                                     |              |      |         |  |  |        |                                                                                                                   |
| Suzuki T., Srivastava A.S., Kurokawa T.;                                               |              |      |         |  |  |        |                                                                                                                   |
| RR "Japanese flounder mRNA for trypsinogen 1."                                         |              |      |         |  |  |        |                                                                                                                   |
| RR Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.                            |              |      |         |  |  |        |                                                                                                                   |
| RC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.     |              |      |         |  |  |        |                                                                                                                   |
| DR EMBL: AB029750; BAA82362.1; -.                                                      |              |      |         |  |  |        |                                                                                                                   |
| DR HSSP; P00763; IDPO.                                                                 |              |      |         |  |  |        |                                                                                                                   |
| DR MEROPS; S01.125;                                                                    |              |      |         |  |  |        |                                                                                                                   |
| DR InterPro; IPR001314; Chymotrypsin.                                                  |              |      |         |  |  |        |                                                                                                                   |
| DR InterPro; IPR001254; Ser_protease_Try.                                              |              |      |         |  |  |        |                                                                                                                   |
| DR PRINTS; PR00122; Chymotrypsin.                                                      |              |      |         |  |  |        |                                                                                                                   |
| DR SMART; SM00020; Tryp_SPC; 1.                                                        |              |      |         |  |  |        |                                                                                                                   |
| DR PROSITE; PS50240; TRYPSIN_DOMAIN; 1.                                                |              |      |         |  |  |        |                                                                                                                   |
| DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.                                           |              |      |         |  |  |        |                                                                                                                   |
| DR PROSITE; PS00135; TRYPSIN_SER; 1.                                                   |              |      |         |  |  |        |                                                                                                                   |
| DR Hydrolase; Serine_protease.                                                         |              |      |         |  |  |        |                                                                                                                   |
| RR SEQUENCE 242 AA; 26348 MW; 6DA722C80BC194A2 CRC64;                                  |              |      |         |  |  |        |                                                                                                                   |
| RR SEQUENCE 244 AA; 26071 MW; F2B8908085BD062 CRC64;                                   |              |      |         |  |  |        |                                                                                                                   |
| RR SEQUENCE 237 TTMAS 243                                                              |              |      |         |  |  |        |                                                                                                                   |
| RR RESULT 13                                                                           |              |      |         |  |  |        |                                                                                                                   |
| RR Q9W7Q6                                                                              |              |      |         |  |  |        |                                                                                                                   |
| RR PRELIMINARY;                                                                        |              |      |         |  |  |        |                                                                                                                   |
| RR ID 09W7Q6;                                                                          |              |      |         |  |  |        |                                                                                                                   |
| RR AC 09W7Q6;                                                                          |              |      |         |  |  |        |                                                                                                                   |
| RR DT 01-NOV-1999 (TREMBLrel. 12, Created)                                             |              |      |         |  |  |        |                                                                                                                   |
| RR DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)                                |              |      |         |  |  |        |                                                                                                                   |
| RR DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)                              |              |      |         |  |  |        |                                                                                                                   |
| RR DE Trypsinogen 2 (Fragment).                                                        |              |      |         |  |  |        |                                                                                                                   |
| RR OS Paralichthys olivaceus (Flounder).                                               |              |      |         |  |  |        |                                                                                                                   |
| RR OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                |              |      |         |  |  |        |                                                                                                                   |
| RR OC Actinopterygii; Neopterygii; Teleostei; Buteoletosteis; Neoteleosteis;           |              |      |         |  |  |        |                                                                                                                   |
| RR OC Acanthomorpha; Acanthopterygii; Percomorphi; Perciformes;                        |              |      |         |  |  |        |                                                                                                                   |
| RR OC Pleuronectoidei; Paralichthyidae; Paralichthys.                                  |              |      |         |  |  |        |                                                                                                                   |
| RR OX NCBI_TaxID=8255;                                                                 |              |      |         |  |  |        |                                                                                                                   |
| RR RN                                                                                  |              |      |         |  |  |        |                                                                                                                   |
| RR RP SEQUENCE FROM N.A.                                                               |              |      |         |  |  |        |                                                                                                                   |
| RR CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY. |              |      |         |  |  |        |                                                                                                                   |
| RR DR EMBL: AB029751; BAA82363.1; -.                                                   |              |      |         |  |  |        |                                                                                                                   |
| RR DR HSPP; P00763; IDPO.                                                              |              |      |         |  |  |        |                                                                                                                   |
| RR DR MEROPE; S01.125;                                                                 |              |      |         |  |  |        |                                                                                                                   |
| RR DR InterPro; IP001314; Chymotrypsin.                                                |              |      |         |  |  |        |                                                                                                                   |
| RR DR InterPro; IP001254; Ser_protease_Try.                                            |              |      |         |  |  |        |                                                                                                                   |
| RR DR Pfam; PF00089; trypsin; 1.                                                       |              |      |         |  |  |        |                                                                                                                   |
| RR DR PRINTS; PRO0722; CHMOTRYSIN.                                                     |              |      |         |  |  |        |                                                                                                                   |
| RR DR SMART; SM00020; TRYSPN_SPC; 1.                                                   |              |      |         |  |  |        |                                                                                                                   |
| RR DR PROSITE; PS50240; TRYPSIN_DOM; 1.                                                |              |      |         |  |  |        |                                                                                                                   |
| RR DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.                                         |              |      |         |  |  |        |                                                                                                                   |
| RR DR PROSITE; PS0135; TRYPSIN_SER; 1.                                                 |              |      |         |  |  |        |                                                                                                                   |
| RR DR HYDRO; SERINE PROTEASE.                                                          |              |      |         |  |  |        |                                                                                                                   |
| RR DR NON_TER                                                                          |              |      |         |  |  |        |                                                                                                                   |
| RR DR SEQUENCE 238 AA; 26071 MW; F2B8908085BD062 CRC64;                                |              |      |         |  |  |        |                                                                                                                   |
| RR SQ                                                                                  |              |      |         |  |  |        |                                                                                                                   |



Oy 179 ITENMICAGFLEGGKDSQGDSGGPVVYCNQELQGIVSWGYGAQKNPKPGVYTKVCNFVDW 238  
| :||||| | :||||| | :||||| | :||||| | :||||| | :||||| | :||||| | :||||| | :||||| |  
Db 180 ITNNMTCIGYLEGGKDSQGDSGGPVVYCNQELQGIVSWGRGCALPNPPGVYTKVCNPNNAW 239  
| :||||| | :||||| | :||||| | :||||| | :||||| | :||||| | :||||| | :||||| | :||||| |  
Oy 239 IQSTIAN 246  
| :|||||  
Db 240 IAQTIAN 247

search completed: June 20, 2003, 15:21:52  
Job time : 85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 20, 2003, 15:06:31 ; Search time 80 Seconds

(without alignments)

411.411 Million cell updates/sec

Title: US-09-762-277a-1

Perfect score: 1327

Sequence: 1 MNPILLIAFLGAAVATPTDD.....VYTKVCNFVDWIQSTIAANS 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:

908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:<sup>\*</sup>

```

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID       | Description                               |
|------------|--------|-------------|--------|-------------|-------------------------------------------|
| 1          | 1327   | 100.0       | 247    | 21 AAY78974 | Canine cationic trypsinogen               |
| 2          | 1111   | 83.7        | 247    | 18 AAW04475 | Porcine trypsinogen                       |
| 3          | 1101   | 83.0        | 246    | 19 AAW64260 | Human amylol beta                         |
| 4          | 1097   | 82.7        | 247    | 23 AAU81693 | Human pancreatic trypsinogen              |
| 5          | 1085   | 81.8        | 247    | 21 AAB21321 | Human trypsinogen                         |
| 6          | 1063.5 | 80.1        | 247    | 21 AAB5701  | Human trypsinogen                         |
| 7          | 1058   | 79.7        | 241    | 21 AAB21316 | Human trypsinogen                         |
| 8          | 1058   | 79.7        | 247    | 9 AAP81243  | Human spleen trypsinogen                  |
| 9          | 1058   | 79.7        | 247    | 16 AAR82703 | Human pancreatic trypsinogen-like protein |
| 10         | 1036   | 78.1        | 240    | 19 AAW57740 |                                           |

## ALIGNMENTS

|                    |                 |                 |                                                     |                                                                                                                                                                      |
|--------------------|-----------------|-----------------|-----------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RESULT 1           | XX              | XX              | XX                                                  | XX                                                                                                                                                                   |
| ID AY78974         | DT 05-JUN-2000  | (first entry)   | DE Canine cationic trypsinogen amino acid sequence. | KW Cationic trypsinogen; dog; monoclonal antibody production; detect; diagnose; pancreatitis; pancreatic cancer; renal insufficiency; extrapancreatic hyposecretion. |
| XX                 | XX              | XX              | XX                                                  | XX                                                                                                                                                                   |
| AC AY78974;        | PR 10-AUG-1998; | PR 10-MAR-1999; | PA (FUJI ) FUJI YAKUHIN KOGYO KK.                   | KW Waritani T, Ashida Y, Yamada T;                                                                                                                                   |
| XX                 | XX              | XX              | XX                                                  | XX                                                                                                                                                                   |
| PN WO200009739-A1. | PD 24-FEB-2000. | PF 09-AUG-1999; | PF 09-AUG-1999;                                     | PR 99WO-JP04299.                                                                                                                                                     |

PT immuno reactants in various forms in diagnosis e.g. of pancreatic diseases  
 XX Claim 3; Page 63-64; 67pp; Japanese.  
 PS This sequence represents the canine cationic trypsinogen amino acid sequence. The protein was isolated from the canine pancreas. The invention relates to monoclonal antibodies with specificity against canine trypsin, or canine trypsin-related substances. The antibodies are highly specific and can be used as reagent for quick and accurate detection and quantitation of canine trypsin and canine trypsin-like immunoreactants in various forms. The antibodies can be used in the diagnosis of diseases such as pancreatitis, pancreatic cancer, renal insufficiency and extrapancreatic hyposecretion.

XX Sequence 247 AA;

Query Match 100.0%; Score 1327; DB 21; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

✓J 1 MNPLILAFGAAVATPDDDKIVGGTCEENSPYQVSINAGTHFCGSSLISDOWVVS 60  
 ✓J 1 MNPLILAFGAAVATPDDDKIVGGTCEENSPYQVSINAGTHFCGSSLISDOWVVS 60  
 QY 61 AHCYKSRIQVRGEYNIDYLEGENEQFINSAKVIRHPNNSWILDNDIMIKLSSPAVLN 120  
 ✓J 61 AHCYKSRIQVRGEYNIDYLEGENEQFINSAKVIRHPNNSWILDNDIMIKLSSPAVLN 120  
 Db 121 ARVATSLPRACAAPGTQCLISGWNTLSSGTNPYPELQCLADAPITLQAQEASYPGQIT 180  
 QY 121 ARVATSLPRACAAPGTQCLISGWNTLSSGTNPYPELQCLADAPITLQAQEASYPGQIT 180  
 Db 121 ARVATSLPRACAAPGTQCLISGWNTLSSGTNPYPELQCLADAPITLQAQEASYPGQIT 180  
 QY 181 ENMICAGFLREGKDSQGDSCGPVYCNQELOGIVSMGYGCAQKNPQGVYTVCNFYDWIQ 240  
 Db 181 ENMICAGFLREGKDSQGDSCGPVYCNQELOGIVSMGYGCAQKNPQGVYTVCNFYDWIQ 240  
 QY 241 STIAANS 247  
 Db 241 STIAANS 247

RESULT<sup>2</sup>  
 AAW08475  
 ID AAW08475 standard; Protein; 247 AA.  
 XX AC AAW08475;  
 XX 28-MAR-1997 (first entry)

DE Porcine trypsinogen.  
 XX Trypsinogen; trypsin; protease; filamentous fungus; Asperillus.  
 XX OS Sus scrofa.  
 PN W09700316-A1.  
 XX PD 03-JAN-1997.  
 XX PF 10-JUN-1996; 96WO-DK00253.  
 XX PR 16-JUN-1995; 95DK-0000693.  
 XX PA (NOVO ) NOVO-NORDISK AS.  
 XX PI Kjeldsen TB, Woldike HF;  
 XX DR WPI; 1997-077523/07.  
 DR N-PSDB; AAT49878.

XX PT Prodn. of trypsin or trypsinogen in filamentous fungi - transformed with recombinant DNA comprising trypsinogen sequence fused to signal

PT sequence, provides higher level of expression than other expression systems  
 XX Example 2; Page 15; 28pp; English.  
 CC An almost complete pig trypsinogen (AAW08475) lacking the very N-terminal end of the signal peptide is the product of a cDNA clone (AAT49878) isolated from a pig pancreas cDNA library. The cDNA clone can be fused to a signal sequence (e.g. from the native enzyme or from a specified amylase gene) and used for prodn. of recombinant trypsin in transformed host cells, esp. filamentous fungus, e.g. Aspergillus oryzae or Aspergillus niger. Levels of trypsin secreted by Aspergillus spp. are several-fold higher than those obt'd. using other microbial systems.

XX Sequence 247 AA;  
 SQ Query Match 83.7%; Score 1111; DB 18; Length 247;  
 Best Local Similarity 81.6%; Pred. No. 4.9e-94;  
 Matches 23; Conservative 23; Mismatches 22; Indels 0; Gaps 0;  
 Matches 200; Conservative 200;  
 QY 2 NPLILAFGAAVATPDDDKIVGGTCEENSPYQVSINAGTHFCGSSLISDOWVVA 61  
 Db 3 NTFFVLLALLGAAGAATPDDDKIVGGTCAANSTPYQVSINSSHFCGSSLISDOWVVA 62  
 QY 62 AHCYKSRIQVRGEYNIDYLEGENEQFINSAKVIRHPNNSWILDNDIMIKLSSPAVLN 121  
 Db 63 AHCYKSRIQVRGEYNIDYLEGENEQFINSAKVIRHPNNSWILDNDIMIKLSSPAVLN 122  
 QY 122 RVATISLPRACAAPGTQCLISGWNTLSSGTNPYPELQCLADAPITLQAQEASYPGQITE 181  
 Db 123 RVATISLPRACAAGTECLISGWNTLSSGTNPYPELQCLADAPITLQAQEASYPGQITG 182  
 QY 182 NMICAGFLREGKDSQGDSCGPVYCNQELOGIVSMGYGCAQKNPQGVYTVCNFYDWIQ 241  
 Db 183 NMICAGFLREGKDSQGDSCGPVYCNQELOGIVSMGYGCAQKNPQGVYTVCNFYDWIQ 242  
 QY 242 TIANN 246  
 Db 243 TIANN 247

RESULT<sup>3</sup>  
 AAW64260  
 ID AAW64260 standard; Protein; 246 AA.  
 XX AC AAW64260;  
 XX DT 24-NOV-1998 (first entry)  
 XX DE Human amyloid beta-protein precursor inhibitor.  
 XX KW MCP-7; mast cell protease 7; trypase-7; serine protease;  
 KW amyloid beta-protein Precursor inhibitor; human; blood clot;  
 KW anticoagulant; myocardial infarction; reocclusion; thromboembolism;  
 KW cerebral embolism; thrombosis; therapy.  
 XX OS Homo sapiens.  
 XX PN WO9824886-A1.  
 XX PR 11-JUN-1998.  
 XX PD 25-NOV-1997; 97WO-US21620.  
 XX PN WO9824886-A1.  
 XX PR 04-DEC-1996; 96US-0032354.  
 XX PA (BIGHM ) BRIGHAM & WOMENS HOSPITAL.  
 XX PI Stevens RL;  
 XX DR WPI; 1998-333008/29.



|                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PN       | JP20000253887-A.                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| PD                    | 14-SEP-2000.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PD       | 19-SEP-2000.                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| PF                    | 09-MAR-2000; 2000WO-CA00258.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PF       | 11-MAR-1999; 99JP-0065337.                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| PR                    | 11-MAR-1999;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| PR                    | 99US-0124260.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | PR       | 11-MAR-1999; 99JP-0065337.                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| PR                    | 01-APR-1999;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| PR                    | 99US-0127386.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | PA       | (TTPH-) TT PHARMA KK.                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| PR                    | 21-JUL-1999;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| XX                    | 99US-0144519.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | XX       | WPI; 2000-558814/64.                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| PA                    | (MOUN ) MOUNT SINAI HOSPITAL.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | DR       | N-PSDB; AAC66182.                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| P1                    | Yousef GM, Diamandis EP;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | PT       | Novel gene encoding a serine protease and its protein used to screen for serine protease inhibitors and to prepare lung disease animal models -                                                                                                                                                                                                                                                                                                                                          |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PT       | PT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| DR                    | 2000-587440/55.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | PS       | Claim 2; Page 3-4; 17PP; Japanese.                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| PT                    | New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L protein mediated disorders, especially cancer.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | CC       | Nucleotide sequence AAC66182 encodes human trypsin hL AAB35701, a serine protease. The invention relates to the human hL gene and protein sequences, to a recombinant vector containing the nucleotide sequence, and a host cell containing the vector. Human trypsin hL can be used for screening for serine protease inhibitors, in the preparation of a lung disease model animal, and for the development of an index marker of lung diseases caused by an anti-trypsin hL antibody. |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | CC       | CC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| PS                    | Example 4; Fig 17: 184pp; English.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | CC       | CC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| PS                    | The present sequence is human trypsinogen, a member of the serine protease family. Kallikreins and kallikrein-like proteins are a subgroup of the serine protease enzyme family. They catalyse the selective cleavage of specific polypeptide precursors to release peptides with potent biological activity. Nucleic acids encoding kallikrein-like proteins KLK-L1, KLK-L2, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins. | CC       | CC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | CC       | CC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| SQ                    | Sequence 247 AA;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | CC       | Sequence 247 AA;                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Query Match           | Score 1085; DB 21; Length 247;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | CC       | Score 1063.5; DB 21; Length 247;                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Best Local Similarity | 80.2%; Pred. No. 1.2e-51;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | CC       | Best Local Similarity 76.1%; Pred. No. 1.1e-89;                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Matches               | 198; Conservative 25; Mismatches 24; Indels 0; Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | CC       | Matches 188; Conservative 31; Mismatches 27; Indels 1; Gaps 1;                                                                                                                                                                                                                                                                                                                                                                                                                           |
| QY                    | 1 MNPLLTIAFLGAAAVATPDDDKIVGGTCBENSVPYQVSLNAGYHFGGSLLSDQWVS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | QY       | 1 MNPLLTIAFLGAAVATPDDDKIVGGTCBENSVPYQVSLNAGYHFGGSLLSDQWVS 59                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Db                    | 1 MNPLLTIAFLGAAVATPDDDKIVGGTCBENSVPYQVSLNAGYHFGGSLLSDQWVS 60                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Db       | 1 MNPLLTIAFLGAAVATPDDDKIVGGTCBENSVPYQVSLNAGYHFGGSLLSDQWVS 60                                                                                                                                                                                                                                                                                                                                                                                                                             |
| QY                    | 61 AAHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 119                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | QY       | 60 SAHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 119                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Db                    | 61 AAHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | Db       | 61 SAHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKTAATL 120                                                                                                                                                                                                                                                                                                                                                                                                                          |
| QY                    | 61 AGHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | QY       | 61 AGHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 120                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Db                    | 61 AGHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 120                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | Db       | 61 AGHCYKSRIQVRLGEYNIDVLEGENQFINSAKVIRHPNYNSWILDNDIMLKSSPAVL 120                                                                                                                                                                                                                                                                                                                                                                                                                         |
| QY                    | 121 ARVATISLPRACAAPTQCLISWGTNTLSSGTNPYELQCLDAPILTQAACEASYPGQIT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | QY       | 120 NAVATISLPRACAAPTQCLISWGTNTLSSGTNPYELQCLDAPILTQAOCESASYPGQI 179                                                                                                                                                                                                                                                                                                                                                                                                                       |
| ?                     | 121 ARVATISLPTAPPATSTKCLISWGWTNTASSGADYPPELQCLDAPIVSLQAKCEASYPGKIT 180                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Db       | 121 NSRVSTVALPRPSCPASAGRCLIVSGWNTLSSGTNPYELQCLDAPIVSLQAKCEASYPGKIT 180                                                                                                                                                                                                                                                                                                                                                                                                                   |
| QY                    | 181 ENMICAGFLEGKGDSCQDGGPVCNGELOGIYSWGYGCAGDNKPGYTKCNGFVDTI 239                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | QY       | 180 ENMICAGFLEGKGDSCQDGGPVCNGELOGIYSWGYGCAGDNKPGYTKCNGFVDTI 239                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Db                    | 181 ENMICAGFLEGKGDSCQDGGPVCNGELOGIYSWGYGCAGDNKPGYTKCNGFVDTI 240                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Db       | 181 ENMICAGFLEGKGDSCQDGGPVCNGELOGIYSWGYGCAGDNKPGYTKCNGFVDTI 240                                                                                                                                                                                                                                                                                                                                                                                                                          |
| QY                    | 241 STIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | QY       | 241 QRTIAAN 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Db                    | 241 NTIAANS 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Db       | 241 QRTIAAN 247                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| RESULT                | 7                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | RESULT   | 6                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| AAB21316              | ID AAB21316 standard; Protein: 241 AA.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | AAB21316 | AAB35701                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| AC                    | AC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | AC       | AC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| XX                    | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | XX       | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| DE                    | Human trypsinogen.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | DE       | Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; trypsinogen; kallikrein-like protein; serine protease; cytosolic; cancer; prostate cancer.                                                                                                                                                                                                                                                                                                                                        |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| KW                    | Homo sapiens.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | KW       | Human; trypsin hL; serine protease; lung disease model animal.                                                                                                                                                                                                                                                                                                                                                                                                                           |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| OS                    | Homo sapiens.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | OS       | W0200053776-A2.                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| XX                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | XX       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| PD                    | 14-SEP-2000.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | PD       | 14-SEP-2000.                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

|                               |                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                        |                    |                                                                        |
|-------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|--------------------|------------------------------------------------------------------------|
| XX                            | PF                                                                       | FT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | FT                                                                     | /label=deletion    |                                                                        |
| XX                            | PF                                                                       | FT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | FT                                                                     | /note="optional"   |                                                                        |
| XX                            | PR                                                                       | FT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | FT                                                                     | /label=deletion    |                                                                        |
| PR                            | 01-MAR-1999;                                                             | 99US-0124260.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | XX                                                                     | /note="optional"   |                                                                        |
| PR                            | 01-APR-1999;                                                             | 99US-0127386.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | PN                                                                     |                    |                                                                        |
| PR                            | 21-JUL-1999;                                                             | 99US-0144919.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | JP63160582-A.                                                          |                    |                                                                        |
| AX                            | PA                                                                       | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | XX                                                                     |                    |                                                                        |
| (MOUN ) MOUNT SINAI HOSPITAL. |                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                        |                    |                                                                        |
| XX                            | PI                                                                       | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | PD                                                                     | 04-JUL-1988.       |                                                                        |
| XX                            | Yousef GM,                                                               | Diamondis EP;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | XX                                                                     |                    |                                                                        |
| XX                            | DR                                                                       | WPI; 2000-587440/55.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | PF                                                                     | 25-DEC-1986;       |                                                                        |
| XX                            | PT                                                                       | New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L protein mediated disorders, especially cancer. -                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | XX                                                                     | 86JP-0307770.      |                                                                        |
| XX                            | PS                                                                       | Example 3; Fig 9; 184pp; English.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | PR                                                                     | 25-DEC-1986;       |                                                                        |
| XX                            | CC                                                                       | The present sequence is human trypsinogen, a member of the kallikrein multi-gene family. Kallikrein and kallikrein-like proteins are a subgroup of the serine protease enzyme family. They catalyse the selective cleavage of specific polypeptide precursors to release peptides with potent biological activity. Nucleic acids encoding kallikrein-like Proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins. | XX                                                                     | 86JP-0307770.      |                                                                        |
| XX                            | Sequence                                                                 | Sequence 241 AA;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | PA                                                                     | (SANY ) SANKYO KK. |                                                                        |
| XX                            | Query Match                                                              | Score 79.7%; Pred. No. 3..5e-89;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | XX                                                                     | XX                 |                                                                        |
| Best Local Similarity         | 79.7%;                                                                   | DB 21; Length 241;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Sequence                                                               | 247 AA;            |                                                                        |
| Matches                       | 192;                                                                     | 25; Mismatches 24;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Score                                                                  | 247;               |                                                                        |
| Conservative                  | 25;                                                                      | Indels 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Match                                                                  | Length 247;        |                                                                        |
| Db                            | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60           | Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | Best Local Similarity                                                  | 79.7%;             |                                                                        |
| Qy                            | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60           | DB 9;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Score                                                                  | 1058;              |                                                                        |
| Db                            | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60           | Pred. No. 3..6e-89;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Mismatches                                                             | 27;                |                                                                        |
| Qy                            | 1 AAHCYKSRIQYRLGEYNIDLEGNEQFINSAKVIRHPNYSWGLDNLIMLKLSSPAVLN 120          | DB 26;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Indels                                                                 | 0;                 |                                                                        |
| Db                            | 1 AAHCYKSRIQYRLGEYNIDLEGNEQFINSAKVIRHPNYSWGLDNLIMLKLSSPAVLN 120          | Gaps 0;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | DB 26;                                                                 | Indels 0;          |                                                                        |
| Qy                            | 1 61 AGHCYKSRIQYRLGEHNIEVLEGNEQFINTNAAKIIRHPQYDRKTLLNIMLKLSSRAVIN 120    | Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60         | Qy                 | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60         |
| Db                            | 1 61 AGHCYKSRIQYRLGEHNIEVLEGNEQFINTNAAKIIRHPQYDRKTLLNIMLKLSSRAVIN 120    | Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60         | Db                 | 1 MNPLLIAPFLGAAVATPDDDKIVGGYCEENSYPOVSILNAGYHFCGGSLISDQWVVS 60         |
| Qy                            | - 121 ARVATISLPRACAAAPGTQCLISGWNNTLSSGNTNPELLOQCLADAPILTQAQCEASYPQQT 180 | Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 61 AAHCYKSRIQYRLGEYNIDLEGNEQFINSAKVIRHPNYSWGLDNLIMLKLSSPAVLN 120       | Qy                 | 61 AAHCYKSRIQYRLGEYNIDLEGNEQFINSAKVIRHPNYSWGLDNLIMLKLSSPAVLN 120       |
| Db                            | - 121 ARVATISLPRACAAAPGTQCLISGWNNTLSSGNTNPELLOQCLADAPILTQAQCEASYPQQT 180 | Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 61 AAHCYKSRIQYRLGEYNIDLEGNEQFINSAKVIRHPNYSWGLDNLIMLKLSSPAVLN 120       | Db                 | 61 AAHCYKSRIQYRLGEYNIDLEGNEQFINSAKVIRHPNYSWGLDNLIMLKLSSPAVLN 120       |
| Qy                            | - 181 ENMICAGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240     | Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 121 ARVATISLPRACAAAPGTQCLISGWNNTLSSGNTNPELLOQCLADAPILTQAQCEASYPQQT 180 | Qy                 | 121 ARVATISLPRACAAAPGTQCLISGWNNTLSSGNTNPELLOQCLADAPILTQAQCEASYPQQT 180 |
| Db                            | - 181 ENMICAGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240     | Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 121 ARVATISLPRACAAAPGTQCLISGWNNTLSSGNTNPELLOQCLADAPILTQAQCEASYPQQT 180 | Db                 | 121 ARVATISLPRACAAAPGTQCLISGWNNTLSSGNTNPELLOQCLADAPILTQAQCEASYPQQT 180 |
| Qy                            | 181 SMNFCVGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240       | Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 181 ENMICAGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240     | Qy                 | 181 ENMICAGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240     |
| Db                            | 181 SMNFCVGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240       | Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 181 SMNFCVGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240     | Db                 | 181 SMNFCVGFLEGGKDSCGDGGPQVVCNGELQGIVSWGYGCAQKPKGVYTKVCNFVDWIQ 240     |
| Qy                            | 241 S 241                                                                | Qy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 241 S 241                                                              | Qy                 | 241 S 241                                                              |
| Db                            | 241 N 241                                                                | Db                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 241 N 241                                                              | Db                 | 241 N 241                                                              |
| XX                            | RESULT 8                                                                 | RESULT 9                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | RESULT 9                                                               | RESULT 9           | RESULT 9                                                               |
| TD                            | AAP81243 standard; protein; 247 AA.                                      | TD                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | AAR82703                                                               | TD                 | AAR82703 standard; Protein; 247 AA.                                    |
| AC                            | AAP81243;                                                                | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | XX                                                                     | XX                 | XX                                                                     |
| XX                            |                                                                          | AC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | AC                                                                     | AC                 | AC                                                                     |
| DE                            | 07-NOV-1990 (first entry)                                                | DE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | DE                                                                     | DE                 | DE                                                                     |
| XX                            | Human spleen trypsin III (trypsinogen III).                              | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | XX                                                                     | XX                 | XX                                                                     |
| KW                            | Human spleen plasminogen; trauma lesions.                                | KW                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | KW                                                                     | KW                 | KW                                                                     |
| XX                            | Homo sapiens.                                                            | OS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | OS                                                                     | OS                 | OS                                                                     |
| FH                            | Location/Qualifiers                                                      | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | XX                                                                     | XX                 | XX                                                                     |
| FT                            | Misc-difference                                                          | XX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | XX                                                                     | XX                 | XX                                                                     |



PT Autocatalytically cleavable zymogenic protease precursors - useful  
PR for cleaving fusion proteins and for therapeutic uses  
PS Disclosure; Fig 1; 45pp; German.  
XX

This invention describes a method where autocatalytically cleavable, zymogenic precursors of a protease (in this invention trypsinogen), have their natural cleavage site replaced by a non natural, autocatalytically cleavable site. Such proteases are reagents for producing peptide hormones and other therapeutic proteins by cleavage of their fusion proteins, and some also have therapeutic activity themselves, e.g. thrombin for treatment of coagulation disorders. The proteases are produced simply and in high yield without the need to add another protease for cleavage, reducing costs and avoiding contamination of the final enzyme.

XX

Sequence 230 AA;

Query Match 77.3%; Score 1026; DB 20; Length 230;  
Best Local Similarity 81.2%; Pred. No. 2..9e-86; Indels 0; Gaps 0;  
Matches 186; Conservative 23; Mismatches 185; Gaps 0;

Qy 19 DDDDKIVGGTCEENSPYQVSINAGHFCGSSLISDOWVVAACHYKSRIQVRGLGYNI 78  
Db 2 DDDDKIVGGTCEENSPYQVSINAGHFCGSSLISDOWVVAACHYKSRIQVRGLGYNI 61

Qy 79 DVLEGNEOFINSAKVIRHPNYSWILDIMIJKLSSPAVLNARVATISILPRAAEGTQ 138  
Db 62 EVLEGNEQFINAAKIIHPQYDRTKLNDIMIJKLSSRAVINARVSTISLTPAPPAGTK 121

Qy 139 CLISRGNTLSSGTNTPELLOQCLADPLLTQACEASPPQTENMICAGFLGGKDSQG 198  
Db 122 CLISRGNTTASSGADYYDELQCLADPLTISQARCEASPPKTISNMPCVTDQ 181

Qy 199 DSGGPVVCNGELQGIVSWGIGCAQKNKPGVTVKCNFVDWIQSTIAANS 247  
Db 182 DSGGPVVCNGQLOGVSVWGDGAQKNKPGVTVKVNWKWIKNTIANS 230

RESULT 12

ID AAY78975 standard; protein; 246 AA.

XX

AC AAY78975;

XX

DT 05-JUN-2000 (first entry)

XX

Canine anionic trypsinogen amino acid sequence.

XX

Anionic trypsinogen; dog; monoclonal antibody production; detect; diagnose; pancreatic cancer; renal insufficiency; extrapancreatic hyposecretion. Canis familiaris.

XX

WO200009739-A1.

XX

FD 24-FEB-2000.

XX

PF 09-AUG-1999; 99WO-JP04229.

XX

PR 10-AUG-1998; 98JP-0236699.

XX

PR 10-MAR-1999; 99JP-006390.

XX

(FUJY ) FUJI YAKUHIN KOGYO KK.

XX

PI Waritani T, Asihida Y, Yamada T;

XX

DR WPI; 2000-206018/18.

XX

Anti-canine trypsin monoclonal antibody, useful as reagent for quick and accurate detection and quantitation of trypsin and/or trypsin-like immunoreactants in various forms in diagnosis e.g. of pancreatic

PT diseases -  
XX Claim 3; Page 64-65; 67pp; Japanese.  
PS XX  
CC This sequence represents the canine anionic trypsinogen amino acid sequence. The protein was isolated from the canine pancreas. The invention relates to monoclonal antibodies with specificity against canine trypsin, or canine trypsin related substances. The antibodies are highly specific and can be used as reagent for quick and accurate detection and quantitation of canine trypsin and canine trypsin-like immunoreactants in various forms. The antibodies can be used in the diagnosis of diseases such as pancreatic cancer, pancreatic cancer, renal insufficiency and extrapancreatic hyposecretion.

XX

Sequence 246 AA;

Query Match 77.3%; Score 1026; DB 21; Length 246;  
Best Local Similarity 75.2%; Pred. No. 3..1e-86;  
Matches 185; Conservative 23; Mismatches 38; Indels 0; Gaps 0;  
Qy 1 MNPLILATLGAYATPTDDDKIVGGYCEBENSVPYQVSINAGHFCGSSLISDOWVVS 60  
Db 1 MKTFIFLAGTAVAFPIDDDKIVGGTCRSRNSPVQVSINAGHFCGSSLISDOWVVS 60

Qy 61 AAHCYKSRIQVRGLGYNIODLEGNEQFINSAKVIRHPNYSWILDIMIJKLSSPAVLN 120  
Db 61 AAHCYKSRIQVRGLGYNIODLEGNEQFINSAKVIRHPNYSWILDIMIJKLSSPAVLN 120

Qy 62 ARVATISIPRACAAPGTOCLISGMNTLSSGTNTPELLOQCLADPLTOAQCEASYPGQIT 180  
Db 121 SRVSAIALPKRSPCAAGTQCLISGMNTQSIGQNPDPDVLOCLAKPLISDSVCRANYPGQIS 180

Qy 181 ENMICGPLEGGSKSCQGDSGGPVCNGELOGQTSWYGIGAQKNKPGVTVKCNFVDWIQ 240  
Db 181 SNMNLGTMEEGGKDSGGPVCNGELOGQTSWYGIGAQKNKPGVSPRKVTVSYWIQ 240

Qy 241 STIAAN 246  
Db 241 QTIRAN 246

RESULT 13

ID AAB98503 standard; Protein; 225 AA.  
AC AAB98503;  
XX AC  
XX DT 03-AUG-2001 (first entry)  
XX DE Human trypsin serine protease catalytic domain.  
KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;  
KW tumour antigen-derived gene 15; serine protease.  
OS Homo sapiens.  
XX PN WO200120056-A1.  
XX PD 26-APR-2001.  
XX PF 20-OCT-2000; 2000WO-US29095.  
XX PR 20-OCT-1999; 990US-0421213.  
XX PA (UYAR-) UNIV ARKANSAS.  
XX PI O'Brien TJ, Tanimoto H;  
XX DR WPI; 2001-3803140.

XX PT Novel extracellular serine protease, termed tumor antigen-derived gene 15 protein overexpressed in carcinomas and DNA encoding it, for PT diagnosis, treatment, prevention of cancer, particularly breast,

PT ovarian cancer  
 XX Example 10; Fig 1; 130PP; English.  
 XX  
 CC The present invention relates to human tumour antigen-derived gene 15  
 CC (TADG-15) protein and coding sequence (see AAH23601 and AAB85001).  
 CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is  
 CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of  
 CC 9-20 residues that lack TADG-15 protease activity are useful for  
 CC vaccinating an individual against TADG-15, having suspected of having or  
 CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a  
 CC diagnostic or therapeutic target in cancer. The present sequence was used  
 CC in a sequence homology alignment with the catalytic domain of TADG-15.  
 XX Sequence 225 AA:  
 23 KIVGGTYCEENSVPYQVSLSNAGYHFCCSLLSDQWVSSAAHCYKSRIQVRLGEYNIDVLE 82  
 DB 1 KIVGGTYCEENSVPYQVSLSNAGYHFCCSLLSDQWVSSAAHCYKSRIQVRLGEHNTEYLE 60

QY 83 GNEQFINSAKVIRHPNNWSWLNDIMLKLSSPAVLNARYATISLSPRACAAPGTCILIS 142  
 DB 61 GNEQFINSAKVIRHPNNWSWLNDIMLKLSSPAVLNARYATISLSPRACAAPGTCILIS 120  
 QY 143 GWGNTLISGTNPYNEPLLQCLDAPILTQAECAASYPGQTENNICAQFLGGKDSQGDSGG 202  
 DB 121 GWGNTLISGTNPYNEPLLQCLDAPILTQAECAASYPGQTENNICAQFLGGKDSQGDSGG 180  
 QY 203 PVVCNGELOGIVSMGYGCAOKNPKGYTKVNFVDWIOSTIAANS 247  
 DB 181 PVVCNGQDQGVSWGDCAQRNPKGYTKVNYVKNNIKNTIAANS 225

Query Match 75.5%; Score 1002; DB 22; Length 225;  
 Best Local Similarity 80.9%; Pred. No. 4.5e-84;  
 Matches 182; Conservative 23; Mismatches 20; Indels 0; Gaps 0;  
 SQ Sequence 224 AA:

24 IVGGTYCENSVPYQVSLSNAGYHFCCSLLSDQWVSSAAHCYKSRIQVRLGEYNIDVLE 83  
 QY 1 IVGGYCEENSVPYQVSLSNAGYHFCCSLLSDQWVSSAAHCYKSRIQVRLGEHNTEYLE 60  
 DB 84 NEQFINSAKVIRHPNNWSWLNDIMLKLSSPAVLNARYATISLSPRACAAPGTCILIS 143  
 DB 61 NEQFINSAKVIRHPNNWSWLNDIMLKLSSPAVLNARYATISLSPRACAAPGTCILIS 120  
 QY 144 WGNTLISGTNPYNEPLLQCLDAPILTQAECAASYPGQTENNICAQFLGGKDSQGDSGG 203  
 DB 121 WGNTLISGTNPYNEPLLQCLDAPILTQAECAASYPGKTISMFCVGPELEGKDSQGDSGG 180

QY 204 VVCGNLOGIVSMGYGCAOKNPKGYTKVNFVDWIOSTIAANS 247  
 DB 181 VVCGNLOGIVSMGYGCAOKNPKGYTKVNFVDWIOSTIAANS 224

RESULT 15  
 ID AAY77494 standard; Protein; 229 AA.  
 XX  
 AC AAY77494;  
 XX  
 DT 05-JUN-2000 (first entry)  
 DE Human trypsin protein domain.  
 DE  
 VW Macrophage stimulating protein; MSP; human; modulator; proliferation;  
 V differentiation; intestinal epithelium; colon crypt; treatment; cancer;  
 W haematopoietic disorder; megakaryocyte deficiency; gastrointestinal;  
 KW chemotherapeutic agent; gut toxicity; serine protease; trypsin.  
 XX  
 OS Homo sapiens.  
 PN US5948892-A.  
 PD 07-SEP-1999.  
 XX  
 PF 16-DEC-1996; 96US-0766992.  
 XX  
 PR 16-DEC-1996; 96US-0766992.  
 PA (AMGE ) AMGEN INC.  
 XX  
 PI Wahl RC;  
 XX  
 DR WPI: 1999-517975/43.  
 XX  
 PT Analogue of macrophage stimulating protein for treating  
 PT gastrointestinal or haematopoietic disorders  
 XX  
 PS Example 2; Column 27-30; 23PP; English.  
 XX

XX  
 CC This invention describes a novel purified and isolated analogue of mature  
 CC macrophage stimulating protein (MSP) having at least one unpaired  
 CC cysteine residue substituted with another amino acid which modulates the  
 CC proliferation or differentiation of the intestinal epithelium. The  
 CC product of the invention binds to RON (a cell membrane protein tyrosine  
 kinase which is a member of the c-met family) to promote the formation of  
 CC colon crypts. MSP analogues are useful for the treatment of conditions  
 CC requiring the administration of MSP, such conditions include  
 CC haematopoietic disorders such as those involving a deficiency of  
 CC megakaryocytes and gastrointestinal disorders such as ulcerative colitis,  
 CC Crohn's disease and infections. The MSP analogues are useful for  
 CC maintaining and repairing the epithelial lining in the treatment of  
 CC cancer, where the aggressive use of chemotherapeutic agents or the use of  
 CC whole body radiation may lead to gut toxicity. The MSP analogues, which  
 CC have a higher activity than normal human MSP are effective at smaller  
 CC dosages. Or optionally, they may be administered less frequently than  
 CC human MSP. This sequence represents a human trypsin serine protease  
 CC domain which is used in a description of the method of the invention.

DR WPI; 2000-195105/17.  
 DR N-PSDB; AAZ2585.

XX Transgenic plant containing DNA encoding a heterologous protein, used  
 XX for the commercial production of protease

PS Example 1: Page 28-29; 32pp; English.

XX The invention relates to a transgenic plant that comprises DNA encoding  
 CC an active or inactive form of a heterologous protease (e.g., trypsin)  
 CC the nucleotide sequence encoding the protease is operably linked to a  
 CC promoter to effect expression of the protease in the plant. The  
 CC transgenic plant is used in a method of producing protease in commercial  
 CC quantities, comprising the extraction of the protease from the plant  
 CC biomass. The current source of trypsin is the organs of animals. One  
 CC of the main difficulties with this source is that there is considerable  
 CC contamination by other proteases. Chymotrypsin is one of the additional  
 CC contaminants, and may leave the product in an undesired manner.  
 CC The method of the invention overcomes the problem of contamination by  
 CC other additional proteases. By directing the expression of the protease  
 CC to the seed of the plant, the problem of expressing the protease and  
 CC having it digest the proteins in the plant, resulting in cell death is  
 CC overcome. The present sequence represents bovine trypsinogen. cDNA  
 CC (Genbank accession #P00760) encoding this protease was used in the  
 CC construction of a plant expression plasmid in the exemplifications of  
 the invention.

XX Sequence 229 AA:

|    | Query Match                                                         | Score | DB                | Length |
|----|---------------------------------------------------------------------|-------|-------------------|--------|
| Qy | Best Local Similarity                                               | 73.7% | 21                | 229;   |
| Db | Matches 176; Conservative                                           | 77.2% | Pred. No. 7.e-82; |        |
| Qy | 29; Mismatches                                                      | 23;   | Indels            | 0;     |
| Db | 0;                                                                  | Gaps  | 0;                |        |
| Qy | 19 DDDDKIVGGTCTEENSPVQVSLSNAGYHFCGGSLISDQMVSAAHCYKSRIQYRGLGEYNI 78  |       |                   |        |
| Db | 2 DDDDKIVGGTCTGANTVPQVSLSNGYHFCGGSLINSQVVSAAHCYKSQIQRGLGEDNI 61     |       |                   |        |
| Qy | 79 DYLEGNEQFINSAKVIRHPNNMSILDIMLKLSSPAVLNARVATISLPRACAAPGTV 138     |       |                   |        |
| Db | 62 NYVEGNEOFISASRSIVPNSNTLNIDMLKLKSASLNSRVASLPLPTSCASAGNIQ 121      |       |                   |        |
| Qy | 139 CLISGWNGLSSGTNYPELQLCDAPITQAQCBAASPYQTENMICAPLEGGKDSCG 198      |       |                   |        |
| Db | 122 CLISGWNGLTKSSGTSPYDVKCLKAPILSDSSCKSAYPGQITSNMFCAGYLEGGKDSCG 181 |       |                   |        |
| Qy | 199 DSGGPVVVCNGELOQIVSNGYCAQRNKPGVYTKVCFNFIQSPIAN 246               |       |                   |        |
| Db | 182 DSGGPVVCSGKLQGIVSNGSGCAOKNPKGVYTKVCNIVSWIKQFIASN 229            |       |                   |        |

Search completed: June 20, 2003, 15:19:52  
 Job time : 82 secs



RESULT 2  
US-08-978-404B-44  
Sequence 44, Application US/08978404B  
~ Patent No. 5968782

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.  
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
FIBRINogen  
NUMBER OF SEQUENCES: 74  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02110-2211  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/978,404B  
FILING DATE: 25-NOV-97  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/032,354  
FILING DATE: 04-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7090  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
TELEX:  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 5968782e

US-08-978-404B-44

Query Match 83.0%; Score 1101; DB 2; Length 246;  
Best Local Similarity 80.5%; Pred. No. 1.5e-106;  
Matches 198; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MNPLILIAFLGAAVATPDDDKIVGGYTCRENSVPYQVSILNAGYHFCGGSLLINDWVVS 60  
Db 1 MSALLILALVGAAVAAFPLEDDDKIVGGTCPEHSVPYQVSILNSGTHFCGGSLLINDWVVS 60.

Qy 61 AAHCYKSRIOVRLGEENIDVLGENSOINSAKVIRIPNYNSWILDNDIMLIKSSPAVLN 120  
Db 61 AAHCYKSRIOVRLGEENIDVLGENSOINSAKVIRIPNYNSWILDNDIMLIKSSPAVLN 120

Qy 121 ARVATISLPRACAAAPCTQCLISWGNTLSSGTNYPELIQLDAPITLQAQCEASYPGQIT 180  
Db 121 ARVAPVAPSACPAQGTDQCLISWGNTLSSGTNYPELIQLDAPITLQAQCEASYPGQIT 180

Qy 181 ENMICAGPLEGGKDSCQGDGGPVVCNGELQGIVSWGNGCAQKNKPGVYTIVCNFVDWIO 240  
Db 181 SMICVGYLEGGRDSCQSDGGPVVCNGOLQGIVSWGIGCALPDNPQVTCNFVGVWIO 240

Qy 241 STAAN 246  
Db 241 DTAAAN 246

RESULT 3  
US-08-467-155A-7  
Sequence 7, Application US/08467155A  
Patent No. 5736377

GENERAL INFORMATION:

APPLICANT: Band, Vimala  
TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED MOLECULES AND METHODS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
APPLICATION NUMBER: US/08/467,155A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00398/100001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-467-155A-7

Query Match 79.8%; Score 1059.5%; DB 1; Length 281;  
Best Local Similarity 72.2%; Pred. No. 3.8e-102;  
Matches 203; Conservative 20; Mismatches 23; Indels 35; Gaps 5;

Qy 1 MNPLILIAFLGAAVATPDDDKIVGGYTCRENSVPYQVSILNAGYHFCGGSLL 53  
Db 1 MSALLILALVGAAVAAFPLEDDDKIVGGTMTRVARTCRESSVPYQVSILNAGYHFCGGSLL 60

Qy 54 SDQWVSSAHCYKSRIOVRLGEENIDVLGENSOINSAKVIRIPNYNSWILDNDIMLIKSSPAVLN 106  
Db 61 NDQWVSSAHCYKRYIQVRLGENIMMTYARINLEGNEQFVTSAKIIRHPNNSWILDNDIMLIKSSPAVLN 120

Qy 107 DMLIKISSPAVL-----NAVATISLPRACAAAPGTQCLISWGNTLSSGTNYPELIQ 159  
Db 121 DMLIKLASPVTLMMTRYARNAYAVSVLPSSCAPAGTQCLISWGNTLSSGTNYPELIQ 180

Qy 160 CLDAPITLQA-----QCEASYPGQITENMICAGFLEGKGDKDSQGDGGPVYCNGLLOG 212  
Db 160 CLDAPITLQA-----QCEASYPGQITENMICAGFLEGKGDKDSQGDGGPVYCNGLLOG 212

Db 181 CVDAPVLPQAMMTRYARDCEASYPGDDTNNMICVGFLEGGRDKDSCGGGGPVCNGELQG 240  
 Qy 213 IVSWGTY-----GCQKQNKPGVYTKVCFNFDWIQSTIAAN 246  
 Db 241 IVSWGTYMTRYARGCAOPDAPGVYTQVTCNYVDWIQNTIADN 281

RESULT 4  
 US-08-628-198-7  
 Sequence 7, Application US/08628198  
 GENERAL INFORMATION:  
 APPLICANT: Band, Vimla  
 TITLE OF INVENTION: NES-1 POLYPEPTIDES, DNA, AND RELATED  
 NUMBER OF SEQUENCES: 11  
 ADDRESS: Fish & Richardson P.C.  
 STREET: 225 Franklin Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/201,038  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/628,198  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/467,155  
 FILING DATE: 06-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Clark, Paul T.  
 REGISTRATION NUMBER: 514  
 TELECOMMUNICATION INFORMATION:  
 REFERENCE/DOCKET NUMBER: 30,162  
 TELEPHONE: 617/542-5070  
 FAX: 617/542-8906  
 TELEPHONE: 617/542-5070  
 TELEFAX: 617/542-8906  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 281 amino acids  
 STRANDEDNESS: not relevant  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-628-198-7

Query Match Score 1059 5; DB 4; Length 281;  
 Best Local Similarity 79.8%; Pred. No. 3.8e-102;  
 Matches 203; Conservative 20; Mismatches 23; Indels 35; Gaps 5;

Qy 1 MNPLILAFQAVATPTDDDKIVGGY-----IDVLEGNEQFINSAKVIRHPNYSWILDN 53  
 Db 1 MSALLILALVGAAVAPVDDDDKIVGGMMTRYARICRESSVPYQVSLNAGHFCCGSLI 60  
 Qy 54 SDQWVSAHCKYKSRIQVRLGEGYN-----IDVLEGNEQFINSAKVIRHPNYSWILDN 106  
 Db 61 NDQWVSAHCKYKIRIQLVRLGEHNMTRYARINVLEGNEQFVDSAKIRHPNYSWILDN 120  
 Qy 107 DMLIKLSSPAVL-----NAVATISLPRACAAPGTQCLISGMNTLSSGTTNPFLQ 159  
 Db 121 DMLIKLSSPAVL-----NAVATISLPRACAAPGTQCLISGMNTLSSGTTNPFLQ 180  
 Qy 160 CLDAPLITQA-----OCEASYPGQTENMICAGFLEGGKDSCOGDGGPVVCNGELQ 212  
 Db 181 CVDAPVLPQAMMTRYARDCEASYPGDTNNMICVGPTEGGKDSCOGDGGPVVCNGELQ 240  
 Qy 213 IVSWGTY-----GCAQKQNKPGVYTKVCFNFDWIQSTIAAN 246  
 Db 241 IVSWGTYMTRYARGCAOPDAPGVYTQVTCNYVDWIQNTIADN 281

Qy 108 CVDAPVLPQAMMTRYARDCEASYPGDTNNMICVGPTEGGKDSCOGDGGPVVCNGELQ 212  
 Db 181 CVDAPVLPQAMMTRYARDCEASYPGDTNNMICVGPTEGGKDSCOGDGGPVVCNGELQ 240  
 Qy 213 IVSWGTY-----GCAQKQNKPGVYTKVCFNFDWIQSTIAAN 246

RESULT 6 - PCT-US96-07343-7

Sequence 7, Application PC/TUS9607343

GENERAL INFORMATION:

APPLICANT: New England Medical Center Hospitals, Inc.

TITLE OF INVENTION: NOVEL POLYPEPTIDES, DNA, AND RELATED MOLECULES AND METHODS

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 245 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/TUS96/07343

FILING DATE: 06-JUN-1995

CLASSIFICATION:

REGISTRATION NUMBER: 30,162

REFERENCE/AGENT INFORMATION:

NAME: Clark, Paul T.

ATTORNEY/AGENT INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 281 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-07343-7

Query Match 79.8%; Score 1059.5; DB 5; Length 281;

~ Best Local Similarity 72.8%; Pred. No. 3.8e-102;

Matches 203; Conservative 20; Mismatches 23; Indels 35; Gaps 5;

Qy 1 MNPLLILAFGLAAVATPDDDKTVTGGY-----TCBENNSVPYQVSLSNAGYHFCGGSLI 53

Db 1 MSALLTLALVGAAVAYFPDDDKTVGGYMMTRYARTCRESSVPYQVSLSNAGYHFCGGSLI 60

Qy 54 SDQWVVAHAHCYKSRIQVRLGEYNN-----NARVATISLPRACAAPGTOCLISGNTLSSGTNPYPELQ 106

Db 61 NDQWVVAHAHCYKVRQQLRGEHNMTRVARYNLPSSCAPAGTQCLISGNTLSSGTNPYPELQ 180

Qy 160 CLDAPILQQA-----OCEASYPGQITENMICAGFLBKGDKDSQGDGGPVYVCNGELQG 212

Db 181 CYDAPVLFQAMTRYARDCEASPGDITNNMICYFLGGKDSQGDGGPVYVCNGELQG 240

Qy 213 IVSWG-----GCAOKNPKPGYTTKVCNFVDTQSTIAAN 246

Db 241 IVSWGMMTRYARGCAQDAPGTYTKVNYVDQTADN 281

Qy Sequence 34, Application US/08944483

Patent No. 6,324,56

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLAAS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.

APPLICANT: STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

TITLE OF INVENTION: OF THE PROSTATE

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cherry L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 224 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6232456e

US-08-944-483-34

Query Match 76.6%; Score 1016; DB 4; Length 224;

~ Best Local Similarity 82.6%; Pred. No. 9.2e-98;

Matches 185; Conservative 20; Mismatches 19; Indels 0; Gaps 0;

Qy 24 IVGGYCEENSVPYQVSLSNAGYHFCGGSLISDQWVVAHAHCYKSRIQVRLGEYNTDVLEG 83

Db 1 IVGGYCEENSVPYQVSLSNAGYHFCGGSLISDQWVVAHAHCYKSRIQVRLGEHNEYLEG 60

Qy 84 NEQFINSAKVTRHPNNWSWLDNDIMLKLSSPAVLNARYATISLPRACAAPGTOCLISG 143

Db 61 NEQFINAAKIRHPKNSRTLDNDLKLSSPAPVINSVSAISLPTAPPAGTESLIG 120

Qy 144 WGNLSSGTNPYPELQCLAPILQTAQCEASYPGQITENMICAGFLGGKDSQGDGGP 203

Db 121 WGNLSSGADYPDQCLAPVLSQAECEASYPGKTNMMFCVGFLGGKDSQGDGGP 180

Qy 204 VVCNEBLQGIVTSWGYCAQDNPKGTYTKVNCFVDTQSTIAANS 247

Db 181 VVSGELQGIVTSWGYCAQDNPKGTYTKVNCFVDTQSTIAANS 224

US-09-027-337-5 Application US/09027337B Sequence 5, Application US/09027337B ; GENERAL INFORMATION: Patent No. 592616 ; APPLICANT: O'Brien, Timothy J. ; APPLICANT: Tanimoto, Hiroto J. ; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in Tissue of Breast and Ovarian Carcinomas ; FILE REFERENCE: D6064 ; CURRENT APPLICATION NUMBER: US/09/027, 337B ; CURRENT FILING DATE: 1998-02-20 ; NUMBER OF SEQ ID NOS: 13 ; SEQ ID NO 5 LENGTH: 225 TYPE: PRT ; ORGANISM: Unknown FEATURE: OTHER INFORMATION: Serine protease catalytic domain of trypsin (TRY) S-09-027-337-5

Query Match 75.5%; Score 1002; DB 2; Length 225; Best Local Similarity 80.9%; Pred. No. 2.6e-96; Matches 182; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

Db 1 KIVGGTCEBENSVPYQVSLSNAGYHFCGGSLISDQWVSAAHCYKSRIQVRGEYNIDVLE 82

Qy 23 KIVGGTCEBENSVPYQVSLSNAGYHFCGGSLISDQWVSAAHCYKSRIQVRGEYNIDVLE 82

Db 1 KIVGGTCEBENSVPYQVSLSNAGYHFCGGSLISDQWVSAAHCYKSRIQVRGEHNIEVE 60

Qy 83 GNEOFINSAKVIRHPNYSWILDNDIMLKLSSPAVLNARVATISLPRACAAPGTOCLIS 142

Db 61 GNEOFINAAKVIIRHPQYDRTKLNDIMLKLSSPAVLNARVATISLPRACAAPGTOCLIS 120

Qy 143 WGNTLSSGTNPPELLQCLAPILTQAQCEASYPQQTENMICAGFLEGKDSCQDSGG 202

Db 121 WGNTASSGADYPDELOCLAPVLSQAKCASYPKITSNMFCVYFLEGGRDSCQDSGG 180

Qy 203 PIVCNGELOGIVSNGYGCQKNPQVYTKVCFNFDWIQSTIAANS 247

Db 181 PIVCNGQLOGGVSMGDGCAQKNPQVYTKVCFNFDWIQSTIAANS 225

RESULT 9 US-09-044-600-5 Sequence 5, Application US/09644600 ; GENERAL INFORMATION: Patent No. 6451500 ; APPLICANT: O'Brien, Timothy J. ; APPLICANT: Tanimoto, Hiroto J. ; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in Carcinomas ; FILE REFERENCE: D6064CIP/D ; CURRENT APPLICATION NUMBER: US/09/644, 600 ; CURRENT FILING DATE: 2000-08-23 ; PRIOR APPLICATION NUMBER: 09/421, 213 ; PRIOR FILING DATE: 1999-10-20 ; PRIOR APPLICATION NUMBER: 09/027, 337 ; PRIOR FILING DATE: 1998-02-20 ; SEQ ID NO 5 LENGTH: 225 TYPE: PRT ; ORGANISM: Homo sapiens FEATURE: OTHER INFORMATION: trypsin US-09-044-600-5

Query Match 75.5%; Score 1002; DB 4; Length 225; Best Local Similarity 80.9%; Pred. No. 2.6e-96; Matches 182; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

Db 1 KIVGGTCEBENSVPYQVSLSNAGYHFCGGSLISDQWVSAAHCYKSRIQVRGEYNIDVLE 82

Qy 83 GNEOFINSAKVIRHPNYSWILDNDIMLKLSSPAVLNARVATISLPRACAAPGTOCLIS 142

Db 61 GNEOFINAAKVIIRHPQYDRTKLNDIMLKLSSPAVLNARVATISLPRACAAPGTOCLIS 120

Qy 143 WGNTLSSGTNPPELLQCLAPILTQAQCEASYPQQTENMICAGFLEGKDSCQDSGG 202

Db 121 WGNTASSGADYPDELOCLAPVLSQAKCASYPKITSNMFCVYFLEGGRDSCQDSGG 180

Qy 144 WGNTLSSGTNPPELLQCLAPILTQAQCEASYPQQTENMICAGFLEGKDSCQDSGG 203

Db 121 WGNTASSGADYPDELOCLAPVLSQAKCASYPKITSNMFCVYFLEGGRDSCQDSGG 180

Qy 204 PIVCNGELOGIVSNGYGCQKNPQVYTKVCFNFDWIQSTIAANS 247

Db 181 PIVCNGQLOGGVSMGDGCAQKNPQVYTKVCFNFDWIQSTIAANS 225

RESULT 11  
US -08 944-483-36  
Sequence 36, Application US/08944483  
Patent No. 6332456

GENERAL INFORMATION:  
 APPLICANT: COHEN, MAURICE  
 APPLICANT: COLETTIS, TRACEY L.  
 APPLICANT: FRIEDMAN, PAULA N.  
 APPLICANT: GRANADOS, EDWARD N.  
 APPLICANT: KLAAS, MICHAEL R.  
 APPLICANT: ROSELL, JOHN C.  
 APPLICANT: STEWART, KENT D.  
 APPLICANT: STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Abbott Laboratories  
 STREET: 100 Abbott Park Road  
 CITY: Abbott Park  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-3500

COMPUTER READABLE FORM:  
 COMPUTER: IBM Compatible  
 MEDIUM TYPE: Diskette  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/944,483  
 FILING DATE:  
 CLASSIFICATION: 4.24  
 PRIORITY APPLICATION DATA:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Becker, Cherry L.  
 REGISTRATION NUMBER: 35,441  
 REFERENCE/DOCKET NUMBER: 6183.US.01  
 TELEPHONE: 847/938-2623  
 TELEFAX: 847/938-2623  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 224 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-08-944-483-36

Query Match 75.1%; Score 997; DB 4; Length 224;  
 Best Local Similarity 80.8%; Pred. No. 8.7e-96;  
 Matches 181; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

RESULT 12  
US -09-296-219-13  
Sequence 13, Application US/09296219  
Patent No. 6248560

GENERAL INFORMATION:  
 APPLICANT: Wahli, Robert C.  
 TITLE OF INVENTION: Analogs of Macrophage Stimulating Protein  
 NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Amgen Inc.  
 STREET: 1840 Dehavilland Drive  
 CITY: Thousand Oaks  
 STATE: California  
 COUNTRY: USA  
 ZIP: 91320-1789

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/296,219  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Winer, Robert B.  
 REFERENCE/DOCKET NUMBER: A-441  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 224 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

US-09-296-219-13

Query Match 75.1%; Score 997; DB 4; Length 224;  
 Best Local Similarity 80.8%; Pred. No. 8.7e-96;  
 Matches 181; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

RESULT 13  
US -08-344-483-35  
Sequence 35, Application US/08944483  
Patent No. 623456

GENERAL INFORMATION:  
 APPLICANT: COHEN, MAURICE  
 APPLICANT: COLETTIS, TRACEY L.  
 APPLICANT: FRIEDMAN, PAULA N.  
 APPLICANT: GRANADOS, EDWARD N.  
 APPLICANT: KLAAS, MICHAEL R.  
 APPLICANT: RUSSELL, JOHN C.  
 APPLICANT: STEWART, KENT D.  
 APPLICANT: STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
 TITLE OF INVENTION: OF THE PROSTATE  
 NUMBER OF SEQUENCES: 76  
 CORRESPONDENCE ADDRESS:  
 ADDRESSE: Abbott Laboratories  
 STREET: 100 Abbott Park Road  
 CITY: Abbott Park  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60064-2500  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/944,483  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Becker, Cheryl L.  
 REGISTRATION NUMBER: 35,441  
 REFERENCE/DOCKET NUMBER: 61,83 .US. 01  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 847/938-2623  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 35:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 224 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: No. 6232456e  
 US-08-944-483-35

Query Match Score 94; DB 4; Length 224;  
 Best Local Similarity 80.4%; Pred. No. 1.8e-55;  
 Matches 180; Conservative 24; Mismatches 20; Indels 0; Gaps 0;

Qy 24 IVGGTCTEENSYPOVSLLNAGYHFCGGLSISPOVWSAAHCYKSRICVRGEYNTDVLEG 83  
 Db 1 IVGGTCTEENSYPOVSLLNAGHFCGGLSISPOVWSAAHCYKTRICVRGPHNIKVLEG 60

-o  
 Qy 84 NEQFINSAKVRHPNNWILDNDIMLKLSSPAVNARATISLPRAAAGTQCLISG 143  
 Db 61 NEQFINAKVRHPKNRDTLNDIMLKLSSPAVNARVSTISLPAPPAGTECLISG 120

Qy 144 WGNTLSSGTNPFLQLDAPLTQACEASYPGQTENMICAGFLEGGDKDSCQGDGGP 203  
 Db 121 WGNTLSSFGADYDPDELKCDLAPLTQACEKASPGKINTSMFCVGFFLEGGDKDSCQRDSGGP 180

Qy 204 YVCNGELQTYSWGCGAQKPGVYTKCNFVDWIGSTIANS 247  
 Db 181 YVCNGQLOQVYSWGCGAWKRNPGVYTKVNYWDWIRDTIANS 224

RESULT 14  
 US-09-120-582-2  
 Sequence 2, Application US/09120582  
 Patent No. 6087558  
 GENERAL INFORMATION:  
 APPLICANT: Howard, John A.  
 TITLE OF INVENTION: COMMERCIAL PRODUCTION OF PROTEASES IN PLANTS  
 FILE REFERENCE: 10014  
 CURRENT APPLICATION NUMBER: US/09/120,582  
 CURRENT FILING DATE: 1998-07-22  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: Patentin Ver. 2.0

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
 TITLE OF INVENTION: OF THE PROSTATE  
 NUMBER OF SEQUENCES: 76  
 CORRESPONDENCE ADDRESS:  
 ADDRESSE: Bos taurus  
 STREET: US-09-120-582-2

Query Match Score 978; DB 3; Length 229;  
 Best Local Similarity 77.2%; Pred. No. 8.5e-94;  
 Matches 176; Conservative 29; Mismatches 23; Indels 0; Gaps 0;

Qy 19 DDDDKIVGGTCTEENSYPOVSLLNAGYHFCGGLSISPOVWSAAHCYKSRICVRGEYNI 78  
 Db 2 DDDDKIVGGTCTCANTVPQVSLLNSGYHFCGGLSISPOVWSAAHCYKSGIQVRGEDNI 61

Qy 79 DVLEGNEOIFNSAKVRHPNNWILDNDIMLKLSSPAVNARATISLPRAAAGTQ 138  
 Db 62 NVVEGNEOITISASSIVSPSYNSNTLNNDIMLKLSSAASLNRSVATISLPTSCASAGTQ 121

Qy 139 CLISGWGNTLSSGTNPFLQLDAPLTQACEASYPGQTENMICAGFLEGGDKDSCQG 198  
 Db 122 CLISGWGNTKSSSTSYSPDYPLKCKAPLSDSSCSAYPOQTSNMFCAGYLEGGDKDSCQG 181

RESULT 15  
 US-08-467-155A-8  
 Sequence 8, Application US/08467155A  
 Patent No. 5736377  
 GENERAL INFORMATION:  
 APPLICANT: Band, Vimala  
 TITLE OF INVENTION: NBS-1 POLYPEPTIDES, DNA, AND RELATED  
 TITLE OF INVENTION: MOLECULES AND METHODS  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 225 Franklin Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 CURRENT APPLICATION DATA:  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 4,335  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Clark, Paul T.  
 REGISTRATION NUMBER: 30,162  
 REFERENCE/DOCKET NUMBER: 00398/100001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/542-3070  
 TELEX: 617/542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 229 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

Query Match Score 950.5; DB 1; Length 299;  
 Best Local Similarity 68.28%; Pred. No. 8.9e-91;  
 Matches 187; Conservative 24; Mismatches 22; Indels 41; Gaps 6;

US-08-467-155A-8

|     |                                                               |                                                                 |     |
|-----|---------------------------------------------------------------|-----------------------------------------------------------------|-----|
| Qy  | 14                                                            | VATPTDDDKIVGGY-----TCCEENSYPOVSLSNAGYHECCGSLISDQWVSAHCY         | 65  |
| Db  | 15                                                            | :     :   :     :     :     :     :     :     :     :     :     | 66  |
| 27  | VAPPFD-DDDKIVGGYHSTRYIVBTS-ENS-PYPOVSLSNSHGECGSLSISEQWVSAHCY  | 85                                                              |     |
| Qy  | 66                                                            | KSRIQYRLGEYN-----IDVLIGNEQFINSKAVIRPHNYNSMILNDIMLKLLSSPA        | 111 |
| Db  | 67                                                            | :     :   :     :     :     :     :     :     :     :     :     | 112 |
| 86  | KTRIQYRLGEHNHSTRYIVBIKVLEGENEQFINAKTRIHPKYNADTLDNDIMLKLLSSPA  | 144                                                             |     |
| Qy  | 118                                                           | VL-----NARATISLPRACAAAGTQCLISMGNTLSGTTTPELLOCDAFDLTOA           | 167 |
| Db  | 119                                                           | :   :     :   :     :     :     :     :     :     :     :     : | 168 |
| 146 | VIHSTRYIVBNARYSTISLPTAPPAGTECLISGNGNTLSFGADYPDELKCLDAFDLTOA   | 201                                                             |     |
| 170 | -----OCEASYPGQTTEINMICAGFLEGKGDSCQGDSSGPVYCNGELOGIVSWGY---    | 218                                                             |     |
| Qy  | 206                                                           | :   :     :   :     :     :     :     :     :     :     :     : | 219 |
| Db  | HSTRYIVBECKASPGKITNSMECVGLFLEGKGDSCQRDSSGPVYCNGLQQGVVSWGHSHST | 265                                                             |     |
| Qy  | 219                                                           | -----GCAOKNKPGYTKICNFWDWQIQTIAANS 247                           | 270 |
| Db  | 226                                                           | RYIBGCAVKNRPGTYTKVNNVWIKDTAANS 299                              | 299 |

Search completed: June 20, 2003, 15:23:14

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 20, 2003, 15:17:26 ; Search time 40 Seconds  
(without alignments)  
593.630 Million cell updates/sec

Title: US-09-762-277A-1  
Perfect score: 1327  
Sequence: 1 MNPLLIAFLGAAVATPTDD.....VYTKVCNFVDWIQSTIAANS 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

total number of hits satisfying chosen parameters:

283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0.8  
Maximum Match 10.0\*

Listing first 45 summaries

Database : PIR\_73;\*

1: Pirl;\*  
2: pir;\*  
3: pir3;\*  
4: Pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description         |
|------------|--------|-------------|--------|----------|---------------------|
| 1          | 1327   | 100.0       | 247    | 1 TRDG   | trypsin (EC 3.4.21) |
| 2          | 1153   | 96.9        | 247    | 2 S13813 | trypsin (EC 3.4.21) |
| 3          | 1127   | 84.9        | 246    | 1 TBRP28 | trypsin (EC 3.4.21) |
| 4          | 1107   | 83.4        | 246    | 1 TRRT2  | trypsin (EC 3.4.21) |
| 5          | 1101   | 83.0        | 246    | 1 TRRT1  | trypsin (EC 3.4.21) |
| 6          | 1098.5 | 82.8        | 247    | 2 A27547 | trypsin (EC 3.4.21) |
| 7          | 1097   | 82.7        | 247    | 1 B25852 | trypsin (EC 3.4.21) |
| 8          | 1085   | 91.8        | 247    | 1 A25852 | trypsin (EC 3.4.21) |
| 9          | 1078   | 80.7        | 231    | 1 TRPGTR | trypsin (EC 3.4.21) |
| 10         | 1058   | 79.7        | 246    | 1 S12764 | trypsin (EC 3.4.21) |
| 11         | 1036   | 78.1        | 304    | 2 S33496 | trypsin (EC 3.4.21) |
| 12         | 1026   | 77.3        | 246    | 1 TRDGC  | trypsin (EC 3.4.21) |
| 13         | 1020.5 | 76.9        | 248    | 2 I38363 | trypsin (EC 3.4.21) |
| 14         | 992.5  | 74.8        | 243    | 2 A35871 | trypsin (EC 3.4.21) |
| 15         | 978    | 73.7        | 229    | 1 TRBOTR | trypsin (EC 3.4.21) |
| 16         | 966.5  | 72.8        | 248    | 2 S55066 | trypsin (EC 3.4.21) |
| 17         | 960.5  | 72.4        | 247    | 2 S05494 | trypsin (EC 3.4.21) |
| 18         | 948.5  | 71.5        | 238    | 2 S31779 | trypsin (EC 3.4.21) |
| 19         | 943    | 71.1        | 248    | 2 S55067 | trypsin (EC 3.4.21) |
| 20         | 905    | 68.2        | 246    | 2 JO1472 | trypsin (EC 3.4.21) |
| 21         | 896    | 67.5        | 246    | 2 S49489 | trypsin (EC 3.4.21) |
| 22         | 877    | 66.1        | 242    | 2 S31775 | trypsin (EC 3.4.21) |
| 23         | 877    | 66.1        | 242    | 2 S31776 | trypsin (EC 3.4.21) |
| 24         | 867    | 65.3        | 231    | 2 S31778 | trypsin (EC 3.4.21) |
| 25         | 855    | 64.5        | 229    | 1 TRDFS  | trypsin (EC 3.4.21) |
| 26         | 854    | 64.4        | 242    | 2 S39048 | trypsin (EC 3.4.21) |
| 27         | 834.5  | 62.9        | 241    | 2 S39047 | trypsin (EC 3.4.21) |
| 28         | 820.5  | 61.8        | 240    | 2 T01779 | trypsin (EC 3.4.21) |
| 29         | 690.5  | 52.0        | 250    |          |                     |

RESULT 2  
S13813

## ALIGNMENTS

|                     |                   |                   |                     |                                         |                                         |                                                  |                     |                                            |                                     |                                                                                              |                                            |                     |                       |                         |                                                                             |                                               |                                                                                                     |                                                                 |                                                                             |                                                           |                                                                           |                                                           |                                                                            |                                                   |                                                                                   |                                                              |                                                                   |                                                                |                                                                |                                                                |                                                                |                                                                |                                                                    |                                                                    |                                                                    |                                                                   |                                                                   |
|---------------------|-------------------|-------------------|---------------------|-----------------------------------------|-----------------------------------------|--------------------------------------------------|---------------------|--------------------------------------------|-------------------------------------|----------------------------------------------------------------------------------------------|--------------------------------------------|---------------------|-----------------------|-------------------------|-----------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------------------|-----------------------------------------------------------|----------------------------------------------------------------------------|---------------------------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------|
| trypsin (EC 3.4.21) | neuropsin - mouse | tissue kallikrein | trypsin (EC 3.4.21) | N;Alternate names: cationic trypsinogen | C;Species: Canis lupus familiaris (dog) | C;Date: 30-Sep-1987 #sequence_change 18-Jun-1999 | C;Accession: A26273 | R;Pinsky, S.D.; LaForge, K.S.; Scheele, G. | Mol. Cell. Biol. 5, 2669-2676, 1985 | A;Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA structure | A;Reference number: A26273; PMID: 3841628; | A;Accession: A26273 | A;Molecule type: mRNA | A;Residues: 1-247 <P1N> | A;Cross-references: GB: M11589; PID: 9164094; PID: AAA30899.1; PID: 9164095 | C;Superfamily: trypsin; homology: trypsinogen | C;Keywords: hydrolase; pancreas; protein digestion; signal sequence; peptide #status predicted <AP> | F:1-154/Domain: signal sequence; peptide #status predicted <SG> | F:16-23/Domain: activation peptide; trypsin, anionic #status predicted <AP> | F:24-239/Domain: trypsin, anionic #status predicted <TRV> | F:30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted | F:63-107,200/Active site: His, Asp, Ser #status predicted | F:75-77,80-85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted | Query Match 100.0%; Score 1327; DB 1; Length 247; | Best Local Similarity 100.0%; Pred. No. 1 2e-105; Mismatches 0; Indels 0; Gaps 0; | Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | QY 1 MNPLLIAFLGAAVATPDDDKIVGGTCEENSPVQYLSNLNAGYHFCGGSLISDQWVVS 60 | 1 MNPLLIAFLGAAVATPDDDKIVGGTCEENSPVQYLSNLNAGYHFCGGSLISDQWVVS 60 | 61 AAHCYKSPRIQVRLGEYNIDVLENEQFTNSAKYTRHPNNSWLNDMLIKSSPAVLN 120 | 61 AAHCYKSPRIQVRLGEYNIDVLENEQFTNSAKYTRHPNNSWLNDMLIKSSPAVLN 120 | 61 AAHCYKSPRIQVRLGEYNIDVLENEQFTNSAKYTRHPNNSWLNDMLIKSSPAVLN 120 | 61 AAHCYKSPRIQVRLGEYNIDVLENEQFTNSAKYTRHPNNSWLNDMLIKSSPAVLN 120 | 121 ARVATISLPRACAPGTOCLISGWNTLSSGTNPYELLOQCLDAPILTORQCEASYPGQT 180 | 121 ARVATISLPRACAPGTOCLISGWNTLSSGTNPYELLOQCLDAPILTORQCEASYPGQT 180 | 121 ARVATISLPRACAPGTOCLISGWNTLSSGTNPYELLOQCLDAPILTORQCEASYPGQT 180 | 181 ENMICAFLEGKDSQGDGGFPVNCNGELOGVSWKGTYCAQNKPGYNTKVCFNEWDMIQ 240 | 181 ENMICAFLEGKDSQGDGGFPVNCNGELOGVSWKGTYCAQNKPGYNTKVCFNEWDMIQ 240 |
|---------------------|-------------------|-------------------|---------------------|-----------------------------------------|-----------------------------------------|--------------------------------------------------|---------------------|--------------------------------------------|-------------------------------------|----------------------------------------------------------------------------------------------|--------------------------------------------|---------------------|-----------------------|-------------------------|-----------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------------------|-----------------------------------------------------------|----------------------------------------------------------------------------|---------------------------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------|

|                                                                                          |  |    |                                                                      |
|------------------------------------------------------------------------------------------|--|----|----------------------------------------------------------------------|
| trypsin (EC 3.4.21.4) - bovine                                                           |  | Qy | 61 AAHCYKSRIOVRLGEYNIDYLEGNEOFINSAKVIRHPPNNSWILDNDIMLIKKLSSPAVLN 120 |
| C;Species: Bos primigenius taurus (cattle)                                               |  | Db | 61 AAHCYKSRIOVRLGEHNINVLEGEOFVDSAKIIRHPNNSWILDNDIMLIKKLSSPAVLN 120   |
| C;Accession: S13813 #sequence_revision 03-aug-1995 #text_change 22-jun-1999              |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| R;le Huerou, I.; Wicker, C.; Guilloteau, P.; Toullée, R.; Puigserver, A.                 |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Eur. J. Biochem. 193, 767-773, 1990                                                      |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic t |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Reference number: S13813; MUID:91065383; PMID:1701147                                  |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Status: preliminary                                                                    |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Molecule type: mRNA                                                                    |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Residues: 1-247 <HUE>                                                                  |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Cross references: EMBL:X54703; NID:9829; PID:CAA38513.1; PID:g830                      |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Superfamily: trypsin; protein homology                                                 |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Keywords: hydrolase; protein digestion; serine proteinase                              |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;24-239/Domain: trypsin homology <TR>                                                   |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;63-107/200/Active site: His, Asp, Ser #status predicted                                |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 86.9%; Score 1153; DB 2; Length 247;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 83.8%; Pred. No. 8.2e-91; Indels 0; Gaps 0;                        |  | Qy | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| Matches 207; Conservative 22; Mismatches 18; Indels 0; Gaps 0;                           |  | Db | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| RESULT 4                                                                                 |  | Qy | 61 AAHCYKSRIOVRLGEYNIDYLEGNEOFINSAKVIRHPPNNSWILDNDIMLIKKLSSPAVLN 120 |
| tryptin (EC 3.4.21.4) II precursor - rat                                                 |  | Db | 61 AAHCYKSRIOVRLGEHNINVLEGEOFVDSAKIIRHPNNSWILDNDIMLIKKLSSPAVLN 120   |
| N;Alternate names: trypsinogen II                                                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Species: Rattus norvegicus (Norway rat)                                                |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 18-Jul-1997              |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Accession: A22657; A00949                                                              |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| R;Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| J. Biol. Chem. 259, 14255-14264, 1984                                                    |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Title: Structure of two related rat pancreatic trypsin genes.                          |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Reference number: A22657; MUID:6094547                                                 |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Accession: A22657                                                                      |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Molecule type: DNA                                                                     |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Residues: 1-246 <CR>                                                                   |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| R;MacDonald, R.J.; Stary, S.J.; Swift, G.H.                                              |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| J. Biol. Chem. 257, 9724-9732, 1982                                                      |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences    |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Reference number: A00948; MUID:82265624; PMID:6896710                                  |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Accession: A00949                                                                      |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Molecule type: mRNA                                                                    |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Residues: 9-246 <MAC>                                                                  |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I    |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Genetics:                                                                              |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Introns: 14/1; 67/2                                                                    |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Superfamily: trypsin; trypsin homology                                                 |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;1-15/Domain: signal sequence #status predicted <API>                                   |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;16-33/Domain: activation peptide #status predicted <API>                               |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;24-246/Product: trypsin II #status predicted <ENZ>                                     |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;24-33/Domain: trypsin homology <TRY>                                                   |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;30-160:48-64:132-233:139-206:171-185/Disulfide bonds: #status predicted                |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;63-107,200/Active site: His, Asp, Ser #status predicted                                |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted               |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 83.4%; Score 1107; DB 1; Length 246;                                         |  | Qy | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| Best Local Similarity 81.3%; Pred. No. 6.7e-87; Indels 0; Gaps 0;                        |  | Db | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| Mismatches 23; Indels 0; Gaps 0;                                                         |  | Qy | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| Matches 200; Conservative 23; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| RESULT 3                                                                                 |  | Qy | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| B225578                                                                                  |  | Db | 1 MNPLLILAFGLAAVATPDDDKIVGGTYCEENSYPPVQYSLNAGYHFCGGSLISDQWVYS 60     |
| trypsin (EC 3.4.21.4) precursor - mouse                                                  |  | Qy | 61 AAHCYKSRIOVRLGEYNIDYLEGNEOFINSAKVIRHPPNNSWILDNDIMLIKKLSSPAVLN 120 |
| C;Species: Mus musculus (house mouse)                                                    |  | Db | 61 AAHCYKSRIOVRLGEHNINVLEGEOFVDSAKIIRHPNNSWILDNDIMLIKKLSSPAVLN 120   |
| C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999              |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| -Accession: B225528                                                                      |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| -Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.                                       |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Nucleic Acids Res. 14, 8307-8330, 1986                                                   |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Title: Sequence organisation and transcriptional regulation of the mouse elastase II a |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Reference number: A93646; MUID:87066713; PMID:3641189                                  |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Accession: B225528                                                                     |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Molecule type: mRNA                                                                    |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| A;Residues: 1-246 <STE>                                                                  |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Superfamily: trypsin; trypsin homology                                                 |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| C;Keywords: calcium binding; hydrolase; protein digestion; serine proteinase             |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;1-23/Domain: signal sequence #status predicted <SIG>                                   |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;24-246/Product: trypsin #status predicted <MAT>                                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;24-33/Domain: trypsin homology <TRY>                                                   |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;30-160:48-64:132-233:139-206:171-185/Disulfide bonds: #status predicted                |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;63-107,200/Active site: His, Asp, Ser #status predicted                                |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted               |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0;                           |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Query Match 84.9%; Score 1127; DB 2; Length 246;                                         |  | Db | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Best Local Similarity 82.5%; Pred. No. 1.3e-88; Indels 0; Gaps 0;                        |  | Qy | 121 ARVATISLPRACAAPTGQCLISWGNTLSSGTNPELLQCLDAPILTOAQCEASYPGQT 180    |
| Matches 203; Conservative 20; Mismatches 23; Indels 0; Gaps 0                            |  |    |                                                                      |

Db 241 DTIAAN 246

A;Title: Isolation and characterization of a cDNA encoding rat cationic trypsinogen.  
 A;Reference number: A27547  
 A;Accession: A27547

RESULT 5

TRRTI

trypsin (EC 3.4.21.4) I precursor - rat

N;Alternate names: trypsinogen I

C;Species: Rattus norvegicus (Norway rat)

C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982

C;Accession: B22657;

A00948

R;Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; Macdonald, R.J.; Rutter, W.J.

J. Biol. Chem. 259, 14255-14264, 1984

A;Title: Structure of two related rat pancreatic trypsin genes.

A;Reference number: A22657; MUID:85054880; PMID:6094547

A;Accession: B22657

A;Molecule type: mRNA

A;Residues: 1-246 &lt;FLE&gt;

A;Cross-references: GB:J00778; PIDN:AAA98518.1; PID:g206508

A;Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 17

A;MacDonald, R.J.; Starý, S.J.; Swift, G.H.

A;Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of

A;Reference number: A00948; MUID:82265624; PMID:6896710

A;Accession: A00948

A;Molecule type: mRNA

A;Residues: 1-246 &lt;NAc&gt;

A;Cross-references: GB:J00778; PIDN:AAA98518.1; PID:g206508

A;Genetics: C;Introns: 14/1; 67/2; 152/1; 197/3

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F;1-15/Domain: signal sequence #status predicted &lt;SIG&gt;

F;16-23/Domain: activation peptide #status predicted &lt;APT&gt;

F;24-246/Domain: trypsin I #status predicted &lt;ENZ&gt;

F;24-239/Domain: trypsin homology &lt;TRY&gt;

F;30-160:48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted

F;63,107/Active site: His, Asp, Ser #status predicted

F;75,77-80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

RESULT 7

B22652

trypsin (EC 3.4.21.4) II precursor [validated] - human

N;Alternate names: trypsin 2; trypsin, antomic; trypsinogen II

C;Species: Homo sapiens (man)

C;Date: 03-Mar-1994 #sequence revision 03-Mar-1994 #text\_change 08-Dec-2000

C;Accession: B22652; A61066; B4988

R;Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Mori, T.; Matsubara, I

Gene 41, 305-310, 1986

A;Title: Immunoreactive anionic and cationic trypsin in human serum.

A;Reference number: A61066; MUID:90091010; PMID:2590466

A;Accession: B22652

A;Molecule type: mRNA

A;Residues: 1-244 &lt;EME&gt;

A;Cross-references: GB:M27602; PID:9521217; PIDN:AAA61232.1; PID:g521218

R;Klimand, M.; Russick, C.; Marks, W.H.; Borgstroem, A.

Clin. Chim. Acta 184, 31-46, 1989

A;Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human trypsinogen.

A;Reference number: A61066; MUID:90091010; PMID:2590466

A;Accession: B22652

A;Molecule type: protein

A;Accession: B43988

A;Experimental source: mucinous ovarian tumor cyst fluid

C;Genetics:

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 22-Jun-1999

C;Accession: A27547

R;Fletcher, T.S.; Alhadef, M.; Craik, C.S.; Largman, C.

Biochemistry 26, 3081-3086, 1987

RESULT 6

A27547

trypsin (EC 3.4.21.4) Precursor, cationic - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 22-Jun-1999

C;Accession: A27547

R;Fletcher, T.S.; Alhadef, M.; Craik, C.S.; Largman, C.

Biochemistry 26, 3081-3086, 1987

A;Map Position: 7q35

A;Note: the human genome contains at least ten trypsin genes or pseudogenes, at least

C;Superfamily: trypsin; trypsin; trypsin homology  
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
 F;1-15/Domain: signal sequence #status predicted <SIG>  
 F;16-24/Domain: activation peptide #status experimental <APT>  
 F;16-24/Domain: trypsinogen II #status experimental <ENZ>  
 F;16-24/Domain: activation peptide #status predicted <APT>  
 F;24-24/Domain: trypsin II #status predicted <TRY>  
 F;24-24/Domain: trypsin homology <TRY>  
 F;63-107/Domain: trypsin homology <TRY>  
 F;63-107/Domain: trypsin homology <TRY>  
 F;63-107/200/Active site: His, Asp, Ser #status predicted  
 F;63-107/200/Active site: His, Asp, Ser #status predicted  
 F;75,77/80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted  
 F;75,77/80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

| Query Match           | Score | Length       | DB 1:    | DB 2:      | DB 3: |
|-----------------------|-------|--------------|----------|------------|-------|
| Best Local Similarity | 81.8% | Pred. No.    | 4.8e-86; |            |       |
| Matches               | 202;  | Conservative | 21;      | Mismatches | 24;   |
| Indels                | 0;    | Gaps         | 0;       |            |       |

Qy 1 MNPLLIAFLGAAVATPTDDDKIVGGTCEENSVPQVSLSNAGYHFCGSSLISDQWVVS 60  
 Db 1 MNPLLITFVAALAAAPFDQQDKIVGGTCEENSVPQVSLSNAGYHFCGSSLISDQWVVS 60  
 Qy 1 MNPLLITFVAAYAAPPFDQQDKIVGGTCEENSVPQVSLSNAGYHFCGSSLISDQWVVS 60  
 ~ 61 AAHCYKSRIOVRLGEYNIDVLEGNEOFINSAKVIRHPNYSWNLNDIMLKSSPVLN 120  
 ~ 61 AAHCYKSRIOVRLGEYNIDVLEGNEOFINSAKVIRHPNYSWNLNDIMLKSSPVLN 120  
 ~ 61 AGHCYKSRIOVRLGEYNIDVLEGNEOFINSAKVIRHPNYSWNLNDIMLKSSPVLN 120  
 ~ 61 ARVATISLPRACAPGTCQCLISWGNTLSSGTNPYELQCLDAPLTQAOCESAYPGKIT 180  
 Qy 121 ARVATISLPRACAPGTCQCLISWGNTLSSGTNPYELQCLDAPLTQAOCESAYPGKIT 180  
 Db 121 ARVATISLPTAPATGPKCLISWGWTASSGADYDPDELQCLDAPLVSQAKCESYPGKIT 180  
 ~ 121 SRVSAILSPPTAPPAGTESLSLGWGNLSSGADYDPDELQCLDAPLVSQAECEASYPGKIT 180  
 Qy 181 ENMTCAGFLGGKDSQGDGGPVYCNQGELQGIVSWGTYGAQKNRPGVTKVNFWDWIK 240  
 Db 181 ENMTCAGFLGGKDSQGDGGPVYCNQGELQGIVSWGTYGAQKNRPGVTKVNFWDWIK 240  
 Qy 241 STIAANS 247  
 Db 241 NTIAANS 247  
 Qy 241 DTIAANS 247  
 Db 241 DTIAANS 247

## RESULT 9

TRPGR

trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)

N;Contains: trypsinogen  
 C;Species: Sus scrofa domestica (domestic pig)  
 C;Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 31-Mar-2000  
 C;Accession: A90641; A90368; A00947  
 R;Charles, M.; Roveri, M.; Guidoni, A.; Desnuelle, P.  
 Biochim. Biophys. Acta 69, 115-129, 1963  
 A;Title: Su le trypsine et la trypsine de porc.  
 A;Reference number: A90641  
 A;Accession: A90641  
 A;Molecule type: protein  
 A;Residues: 1-10 <CHA>  
 R;Hermanson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.  
 Biochemistry 12, 3146-3153, 1973  
 A;Title: Determination of the amino acid sequence of porcine trypsin by sequenator ar  
 A;Accession: A90368; MUID:73228692; PMID:4738933  
 A;Molecule type: protein  
 A;Residues: 9-231 <HFR>  
 A;Note: at position 20, Ile and Val occur alternatively  
 C;Superfamily: trypsin; trypsin homology  
 C;Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase;  
 F;1-8/Domain: trypsinogen #status experimental <ZYM>  
 F;1-8/Domain: activation peptide #status experimental <APT>  
 F;9-23/Domain: trypsin #status experimental <MAT>  
 F;9-22/Domain: trypsin homology <TRY>  
 F;15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted  
 F;48,92,185/Active site: His, Asp, Ser #status predicted  
 F;60,62,65,0/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted  
 Qy 17 PRDDDKIVGGTCEENSVPQVSLSNAGYHFCGSSLISDQWVVSAAHCYKSRIQLVRGEBY 76  
 Db 2 PRDDDKIVGGTCAANSIPIQVSLSNAGYHFCGSSLISDQWVVSAAHCYKSRIQLVRGEBY 61

R;Wiegand, U.; Corbach, S.; Minn, A.; Kang, J.; Mueller-Hill, B.  
submitted to the EMBL Data Library; March 1993  
A;Description: Identification, cloning and characterization of a cDNA encoding a human  
A;Reference number: S33496  
A;Accession: S33496  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-304 <WIB>  
A;Cross-references: EMBL:X72781  
C;Genetics:  
A;Gene: GDB:PRSS3; TRY4  
A;Cross-references: GDB:335300  
A;Map Position: 7q35-7q35  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine protease  
P;81-96/Domain: trypsin homology <TRY>  
F;120-164/257/Active site: His, Asp, Ser #status predicted

**RESULT 10**

trypsin (EC 3.4.21.4) III precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 22-Jun-1999  
,Accession: S12764  
A;Gene: GDB:PRSS3; TRY3  
A;Cross references: GDB:335297  
A;Cross position: 7q35-7q35  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: calcium binding; hydrolase; pancreas; protein digestion; serine proteinase;  
P;16-21/Domain: signal sequence #status predicted <SIG>  
P;22-247/Product: trypsin III #status predicted <APT>  
P;30-160,48-64,139-206,171-185/Disulfide bonds: #status predicted  
P;63,107,200/Active site: His, Asp, Ser #status predicted  
P;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 79.7%; Score 1058; DB 2; Length 247;  
Best Local Similarity 78.5%; Pred. No. 1e-83;  
Matches 194; Conservative 27; Mismatches 26; Indels 0; Gaps 0;

Qy 1 MNPLLIPLAFAVAVPTDDDKTIVGGTCBENSVPYQVSLSNAGTYFCGSSLISDQWVWS 60  
, 1 MNPFILATGAAAVAVPFDDDKTIVGGTCCEENSLPYQVSLSNGSHFCGSSLISDQWVWS 60

Qy 61 AAHCYKSRIOVRLGEYNIDYLEGENEQTINSAKVIRHPPNNWSILDIMIJKLSSPAVLN 120  
Db 61 AAHCYKTRIQVRLGEHNKTVLEGNEQFINAKIRHPKNRDTLDNDIMIJKLSSPAVIN 120

Qy 121 ARVATISLPRACAAPTQCLISGNGNTLSGGTNPYELLQCDAPILTQACEASTYQGQT 180  
Db 121 ARVSTSPPTAPPAGTEUJISGNGNTLSGGTNPYELLQCDAPILTQACEASTYQGQT 180

Qy 181 ENMICAGFLEGGKDSQGDGGPPYVCNGELOQIVSGWYGCQAQNPKPGVYTKCNFVDWIQ 240  
, 1 NMFCVGFLLEGKDSWKRDSGGPVVCNGQGLQGVVSWGHGCAWKNRPGVITKVNVFDWIK 240

Db 181 NMFCVGFLLEGKDSWKRDSGGPVVCNGQGLQGVVSWGHGCAWKNRPGVITKVNVFDWIK 240

Qy 241 STIAANS 247  
Db 241 DTIAANS 247

**RESULT 11**

trypsin (EC 3.4.21.4) IV form a - human  
C;Species: Homo sapiens (man)  
C;Date: 03-Mar-1991 #sequence\_revision 03-Aug-1995 #text\_change 15-Aug-1997  
C;Accession: S33496

Query Match 77.3%; Score 1026; DB 1; Length 246;  
Best Local Similarity 75.2%; Pred. No. 5.3e-80;  
Matches 185; Conservative 23; Mismatches 38; Indels 0; Gaps 0;

Qy 1 MNPLLIPLAFAVAVPTDDDKTIVGGTCBENSVPYQVSLSNAGTYFCGSSLISDQWVWS 60  
Db 1 MKTFIFALLGATVAFFDDDKTIVGGTCBENSVPYQVSLSNAGTYFCGSSLISDQWVWS 60

Qy 61 AAHCYKSRIOVRLGEYNIDYLEGENEQTINSAKVIRHPPNNWSILDIMIJKLSSPAVLN 120



F:6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental  
 F;13-143,31-47,115-168,179-203/Disulfide bonds: #status experimental  
 F;46,90-183/Active site: His, Asp, Ser #status experimental  
 F;58-60,63,68/Binding site: calcium (Glu, Asp, Val, Glu) #status experimental  
 F;131-132/Cleavage site: Lys-ser (autoytic) #status experimental

| Query Match           | Score | Length       | DB 1: | DB 2:            |
|-----------------------|-------|--------------|-------|------------------|
| Best Local Similarity | 73.7% | 29;          | DB 1: | Length 29;       |
| Matches               | 176;  | Conservative | DB 1: | Pred. No. 6e-76; |
|                       |       | Mismatches   | DB 1: | Indels 0;        |
|                       |       |              | DB 1: | Gaps 0;          |

QY 19 DDDDKVGGTCEENSPYQVSLSNAGYHFCGGLSISDQWVSSAHCYKSRIQVRLEYNI 78  
 |:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 2 DDDDKVGGTCCGANTVPYQVSLSNQHFCGGLSINSQWVSSAHCYKSGTQVRLEGDN 61

QY 79 DYLEGNEOFINSAKVIRHPNNWSWILDNDIMLKLSSPAVLNARYATISLPRACAPGTQ 138  
 :|:|||||:||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 62 NVVEGNEQFISASKSTVHPNTNSNTLNDIMLKLRSAAASLNSRVASISLPTSCASAGTQ 121

QY 139 CLISGWGNTLSSGTYNPELLQCLDAPILTQAQCEASYPGQTENMICAGLEGGKDSCG 198  
 |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 .b. 122 CLISGWGNTTSSGTYPDVNLKCLKAPILSUSSCKBAYPGQITSNMFCAGYLEGGKDSCG 181

QY 199 DSGGPVVCNGELOGTVSGYGAQRNKPGYTKVCFWDIQSTIAAN 246  
 |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 182 DSGGPVVCSCGKLGQIVSWGSGCAQRNKPVGVTKVCNVSNIKTIASN 229

Search completed: June 20, 2003, 15:22:34  
 Job time : 41 secs